

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:42:01 ; Search time 3702.44 Seconds
(without alignments)
10441.655 Million cell updates/sec

Title: US-09-671-050-11

Perfect score: 945
Sequence: 1 atggaaaagtatgaaaaatt.....aggtaacttcgctcaaaagt 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1:	gb.ba.*
2:	gb.htg.*
3:	gb.in.*
4:	gb.om.*
5:	gb.ov.*
6:	gb.pat.*
7:	gb.ph.*
8:	gb.pl.*
9:	gb.pr.*
10:	gb.ro.*
11:	gb.sts.*
12:	gb.sy.*
13:	gb.un.*
14:	gb.vi.*
15:	em.ba.*
16:	em.fun.*
17:	em.fun.*
18:	em.in.*
19:	em.mu.*
20:	em.on.*
21:	em.ox.*
22:	em.ov.*
23:	em.pat.*
24:	em.ph.*
25:	em.pl.*
26:	em.ro.*
27:	em.sts.*
28:	em.un.*
29:	em.vi.*
30:	em.htg_hum.*
31:	em.htg_inv.*
32:	em.htg_other.*
33:	em.htg_mus.*
34:	em.htg_pln.*
35:	em.htg_rod.*
36:	em.htg_man.*
37:	em.htg_vrt.*
38:	em.sy.*
39:	em.htgo_hum.*
40:	em.htgo_mus.*
41:	em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	945	100.0	945	6	AX107722	AX107722 Sequence
2	929	98.3	972	6	AX107716	AX107716 Sequence
3	868.8	91.9	1678	6	AX698820	AX698820 Sequence
4	863	91.3	1083	6	AX166534	AX166534 Sequence
5	839	88.8	1041	6	AX107720	AX107720 Sequence
6	823	87.1	1068	6	AX107714	AX107714 Sequence
7	647	68.5	687	6	AX286069	AX286069 Sequence
8	647	68.5	882	6	AX286067	AX286067 Sequence
9	643	68.0	911	6	AX056404	AX056404 Sequence
10	559.2	59.2	2615	6	AX056405	AX056405 Sequence
11	507	53.7	1819	6	AX107724	AX107724 Sequence
12	454.4	48.1	561	6	AX107712	AX107712 Sequence
13	454.4	48.1	594	6	AX107718	AX107718 Sequence
14	415.4	44.0	1612	6	AX399975	AX399975 Sequence
15	413.8	43.8	1363	9	HSSTHPKB	X66358 H.sapiens m
16	400.2	42.3	2952	3	AK112442	AK112442 Ciona int
17	287	30.4	1701	6	AR221995	AR221995 Sequence
18	287	30.4	3080	6	AR221994	AR221994 Sequence
19	287	30.4	3153	4	AB029045	AB029045 Oryctolag
20	283.8	30.0	1993	9	HSU35146	U35146 Homo sapien
21	272.6	28.8	1518	10	AB029067	AB029067 Mus muscu
22	272.6	28.8	3988	10	AB029066	AB029066 Mus muscu
23	272.6	28.8	4022	10	AB029065	AB029065 Mus muscu
24	227.4	24.1	1513	9	AF130372	AF130372 Homo sapi
25	227.4	24.1	1773	6	AX166533	AX166533 Sequence
26	227.4	24.1	2146	9	BC041799	BC041799 Homo sapi
27	222.6	23.6	1620	10	BC028871	BC028871 Mus muscu
28	218.4	23.1	101340	2	AC018104	AC018104 Drosophil
29	218.4	23.1	190668	3	AC007419	AC007419 Drosophil
30	218.4	23.1	260731	3	AE003611	AE003611 Drosophil
31	211.4	22.4	1732	10	AF112183	AF112183 Rattus no
32	211.4	22.4	1738	10	AF112184	AF112184 Rattus no
33	208.4	22.1	1974	3	AF326965	AF326965 Trypanoso
34	200.4	21.2	80362	9	AC079615	AC079615 Homo sapi
35	194	20.5	3399	9	AY217744	AY217744 Homo sapi
36	194	20.5	3399	9	HSTXP3	Y15057 Homo sapien
37	194	20.5	3441	9	BC036091	BC036091 Homo sapi
38	180.2	19.1	3321	3	LM293284	AX293284 Leishmani
39	175.8	18.6	1866	6	AX056407	AX056407 Sequence
40	171.8	18.2	903	8	AF026545	AF026545 Pneumocys
41	171.4	18.1	2982	9	AF390028	AF390028 Homo sapi
42	168	17.8	960	3	TCU69958	U69958 Trypanosoma
43	163.6	17.3	921	3	DDICRP	L00652 Dictyostel
44	163.4	17.3	1476	6	AR087351	AR087351 Sequence
45	163.4	17.3	1476	6	AR281916	AR281916 Sequence

ALIGNMENTS

RESULT 1
AX107722
LOCUS AX107722 945 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 11 from Patent WO01233579.
ACCESSION AX107722
VERSION AX107722.1 GI:139233202
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and
AUTHORS Sande,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same

JOURNAL Patent: WO 0123579-A 11 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES Location/Qualifiers
source 1..945
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 165 c 202 g 255 t
ORIGIN

Query Match 100.0%; Score 945; DB 6; Length 945;
Best Local Similarity 100.0%; Pred. No. 8.6e-206;
Matches 945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAAGAATGATAAATTAGCTTAAGACTGGAGAAGGCTCTTATGGGTTGTATTCAA 60
Db 1 ATGGAAGAATGATAAATTAGCTTAAGACTGGAGAAGGCTCTTATGGGTTGTATTCAA 60

Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTTAAATAATTGTGGAACTCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTTAAATAATTGTGGAACTCTGAAGAT 120

Qy 121 GATCCTGTTGTAAAGAAATAGCACTAAGAGAAATAGTATGTTGAAGCAACTTAAACAT 180
Db 121 GATCCTGTTGTAAAGAAATAGCACTAAGAGAAATAGTATGTTGAAGCAACTTAAACAT 180

Qy 181 CCAATCTTGTGAACCTCATCGAGTGTTTTCAGGAGAAAGGAAATGCAATTTAGTTTTT 240
Db 181 CCAATCTTGTGAACCTCATCGAGTGTTTTCAGGAGAAAGGAAATGCAATTTAGTTTTT 240

Qy 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300

Qy 301 GGAGTGATCAAAAGCGTATTATGCAAACTTCAAGCTCTTAATTTCTGTCTATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGCAAACTTCAAGCTCTTAATTTCTGTCTATATACAT 360

Qy 361 AACTGTATTTCACAGAGATATAAAACCTGAAATAATTTCTAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTTCACAGAGATATAAAACCTGAAATAATTTCTAACTAAGCAAGGAATAATC 420

Qy 421 AAGATTTGTGACTTCGGGTTTTCAGAACTTCAAGCTCTTAATTTCTGTCTATATACAT 480
Db 421 AAGATTTGTGACTTCGGGTTTTCAGAACTTCAAGCTCTTAATTTCTGTCTATATACAT 480

Qy 481 GTAGCTACAGAGATGTTACCGAGCTCTTGAACCTTCTGTGGAGATACCTCAGTATGGTTCT 540
Db 481 GTAGCTACAGAGATGTTACCGAGCTCTTGAACCTTCTGTGGAGATACCTCAGTATGGTTCT 540

Qy 541 TCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCTTTCAGAGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCTTTCAGAGCCAGCCACTG 600

Qy 601 TGGCTCGAAATCAGATGTGGCAACTTATCTGATTAATCAGAACACTAGGAAATTA 660
Db 601 TGGCTCGAAATCAGATGTGGCAACTTATCTGATTAATCAGAACACTAGGAAATTA 660

Qy 661 ATCCAGAGACATCAATCAATCTTTAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCT 720
Db 661 ATCCAGAGACATCAATCAATCTTTAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCT 720

Qy 721 GAGCCAGAGACATGGAACCTTTTCAGGAAAGTTTTCAGATGTTCTCCTGTGGGCTCTG 780
Db 721 GAGCCAGAGACATGGAACCTTTTCAGGAAAGTTTTCAGATGTTCTCCTGTGGGCTCTG 780

Qy 781 AACTTCATGAAGGGGTGCTCAAGATGAATCCAGATGACAGATTAACTGTCTCCCACTC 840
Db 781 AACTTCATGAAGGGGTGCTCAAGATGAATCCAGATGACAGATTAACTGTCTCCCACTC 840

Qy 841 CTGGAGAGCTCCTACTTTGATTTCTTTTCAAGAGGCCCAAAATTAAGAAAGAACGACGTAAT 900
Db 841 CTGGAGAGCTCCTACTTTGATTTCTTTTCAAGAGGCCCAAAATTAAGAAAGAACGACGTAAT 900

Qy 901 GAAGGAAGAAACAGAAAGACGCCAACAGCTACTTCCGCTCAAAAGT 945
Db 901 GAAGGAAGAAACAGAAAGACGCCAACAGCTACTTCCGCTCAAAAGT 945

RESULT 2
AX107716
LOCUS AX107716 972 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 5 from Patent WO0123579.
ACCESSION AX107716
VERSION AX107716.1 GI:13923199
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and
Sands, A.T.
Humar kinase proteins and polynucleotides encoding the same
Patent: WO 0123579-A 5 05-APR-2001;
Lexicon Genetics Incorporated (US);
FEATURES
Location/Qualifiers
1..972
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 172 c 214 g 263 t
ORIGIN

Query Match 98.3%; Score 929; DB 6; Length 972;
Best Local Similarity 100.0%; Pred. No. 3.9e-202;
Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAAGAATGATAAANTAGCTTAAGACTGGAGAAGGCTCTTATGGGTTGTATTCAA 60
Db 1 ATGGAAGAATGATAAANTAGCTTAAGACTGGAGAAGGCTCTTATGGGTTGTATTCAA 60

Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTTAAATAATTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTTAAATAATTGTGGAATCTGAAGAT 120

Qy 121 GATCCTGTTGTAAAGAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACAT 180
Db 121 GATCCTGTTGTAAAGAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACAT 180

Qy 181 CCAATCTTGTGAACCTCATCGAGTGTTTTCAGGAGAAAGGAAATGCAATTTAGTTTTT 240
Db 181 CCAATCTTGTGAACCTCATCGAGTGTTTTCAGGAGAAAGGAAATGCAATTTAGTTTTT 240

Qy 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300

Qy 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACAT 360

Qy 361 AACTGTATTTCACAGAGATATAAAACCTGAAATAATTTCTAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTTCACAGAGATATAAAACCTGAAATAATTTCTAACTAAGCAAGGAATAATC 420

Qy 421 AAGATTTGTGACTTCGGGTTTTCAGAACTTCAAGCTCTTAATTTCTGTCTATATACAT 480
Db 421 AAGATTTGTGACTTCGGGTTTTCAGAACTTCAAGCTCTTAATTTCTGTCTATATACAT 480

Qy 481 GTAGCTACAGAGATGTTACCGAGCTCTTGAACCTTCTGTGGAGATACCTCAGTATGGTTCT 540
Db 481 GTAGCTACAGAGATGTTACCGAGCTCTTGAACCTTCTGTGGAGATACCTCAGTATGGTTCT 540

Qy 541 TCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCTTTCAGAGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCTTTCAGAGCCAGCCACTG 600

```
Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTTCAGAGCTCCTGACAGGCCAGCCACTG 660
Qy 601 TGGCTCGAATATCAGATCTGGACCACTTTTATCTGTAATACAGACACTAGGAAATTA 660
Db 601 TGGCTCGAATATCAGATCTGGACCACTTTTATCTGTAATACAGACACTAGGAAATTA 660
Qy 661 ATCCCAAGACATCAATCAATCTTTAAAAGTAACGGTTTTTCCATGGCATCAGTATACCT 720
Db 661 ATCCCAAGACATCAATCAATCTTTAAAAGTAACGGTTTTTCCATGGCATCAGTATACCT 720
Qy 721 GAGCCAGAGACATGGAACCTCTGAGGAAAGTTCTCAGATGTTTCATCTGGGCTCTG 780
Db 721 GAGCCAGAGACATGGAACCTCTGAGGAAAGTTCTCAGATGTTTCATCTGGGCTCTG 780
Qy 781 AACTTCATGAAGGGGTGTCGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTC 840
Db 781 AACTTCATGAAGGGGTGTCGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTC 840
Qy 841 CTGGAGAGCTCTACTTGTATCTTTTCAAGAGGCCCAATTAAGAAAGACAGTAAT 900
Db 841 CTGGAGAGCTCTACTTGTATCTTTTCAAGAGGCCCAATTAAGAAAGACAGTAAT 900
Qy 901 GAAGGAGAAACAGAGAGCCCAACAGGT 929
Db 901 GAAGGAGAAACAGAGAGCCCAACAGGT 929

RESULT 3
LOCUS AX698820 1678 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 6 from Patent WO02079473.
ACCESSION AX698820
VERSION AX698820.1 GI:29499608
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Panzer,S.R., Lincoln,S.E., Aitus,C.M., Dufour,G.E., Hillman,J.L.,
Jones,A.L., Dam,T.C., Liu,T.F., Harris,B., Flores,V., Daffo,A.,
Marwaha,R., Chen,A.J., Chang,S.C., Gershin,J.E., Peralta,C.H.,
David,M.H. and Lewis,S.A.
TITLE Molecules for diagnostics and therapeutics
JOURNAL Patent: WO 02079473-A 6 10-OCT-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1678
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notes="Incyte ID No: LI:058298.1:2001CAN12"
BASE COUNT 541 a 308 c 336 g 493 t
ORIGIN

Query Match 91.9%; Score 868.8; DB 6; Length 1678;
Best Local Similarity 99.2%; P-val No. 2.2e-188;
Matches 925; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 1 ATGCAAAAGTATGAAATATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATCAAA 60
Db 131 ATGCAAGAGTATGAAATATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATCAAA 190
Qy 61 TCAGAGACACCAACCTCTGGACAGTACTAGCTGTAAATAATTTGGATCTGAGA 119
Db 191 TCAGAGTAACAAACCTCTGGACAGTACTAGCTGTAAATAATTTGGATCTGAGA 250
Qy 120 TGATCTGTGTTTAAAGAAATACACTAAGAGAAATACGTA-TGTTGAAGCAATTAAC 178
Db 251 TGATCTGTGTTTAAAGAAATACACTAAGAGAAATACGTA-TGTTGAAGCAATTAAC 310
Qy 179 ATCAAAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATATGCAATTTAGTTT 238
Db 179 ATCAAAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATATGCAATTTAGTTT 238
```

```
Db 311 ATCAAAATCTTTGAAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATATGCAATTTAGTTT 370
Qy 239 TTGTAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGCTTGGTG 298
Db 371 TTGTAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGCTTGGTG 430
Qy 299 ATGGAGTGTATCAAAAGGCTTATTATGGCAACACTTCAAGCTCTTTAAATTTCTGTCTATAC 358
Db 431 ATGGAGTGTATCAAAAGGCTTATTATGGCAACACTTCAAGCTCTTTAAATTTCTGTCTATAC 490
Qy 359 ATAACCTGTATTCACAGAGATATAAAACCTGAAATATCTTAATTAATTAAGCAAGGATAA 418
Db 491 ATAACCTGTATTCACAGAGATATAAAACCTGAAATATCTTAATTAATTAAGCAAGGATAA 550
Qy 419 TCAAGATTTTGACTTTCGGGTTTGCAAAATTTCTGATTTCCAGGAGATGCTTACACCGATT 478
Db 551 TCAAGATTTTGACTTTCGGGTTTGCAAAATTTCTGATTTCCAGGAGATGCTTACACCGATT 610
Qy 479 ATGTAGCTA-CGAGATGTACCGAGCT-CCTGAACTTCTTGTGGGAGATACT-CAGTATG 535
Db 611 ATGTAGCTAGCGAGATGTACCGAGCTCCTCGAACTTCTTGTGGGAGATACTCCTCAGTATG 670
Qy 536 GTTCTTCAGTCGATATATGGGCTATTGGTCTGTTTGTGAGAGCTCCTGACAGGCCAGC 595
Db 671 GTTCTTCAGTCGATATATGGGCTATTGGTCTGTTTGTGAGAGCTCCTGACAGGCCAGC 730
Qy 596 CACTGTGGCTGGAAATCAGATGTGGACCACTTTATCTGATTAATCAGAACACTAGGAA 655
Db 731 CACTGTGGCTTGAATTCAGATGTGGACCACTTTATCTGATAATCAGAACACTAGGAA 790
Qy 656 AATTAAATCCAGAGATCAATCAATCTTTTAAAGTAAAGGGTTTTTCCATGGCATCAGTA 715
Db 791 AATTAAATCCAGAGATCAATCAATCTTTTAAAGTAAAGGGTTTTTCCATGGCATCAGTA 850
Qy 716 TACCTGAGCCAGAGACATGGAACCTTTGAGGAAAGTTCTCAGATGTTTCATCTCTGG 775
Db 851 TACCTGAGCCAGAGACATGGAACCTTTGAGGAAAGTTCTCAGATGTTTCATCTCTGG 910
Qy 776 CTCTGAACTTCTAAGGGGTGCTCGAAGTGAATCCAGATGACAGATTAACCTGTTCCC 835
Db 911 CTCTGAACTTCTAAGGGGTGCTCGAAGTGAATCCAGATGACAGATTAACCTGTTCCC 970
Qy 836 AACTCTCTGAGAGCTCCTACTTTGATTTGATTTCTTTCAAGAGGCCCAATTAAGAAAGAAAGCAC 895
Db 971 AACTCTCTGAGAGCTCCTACTTTGATTTGATTTCTTTCAAGAGGCCCAATTAAGAAAGAAAGCAC 1030
Qy 896 GTAATGAAGAAAGAAACAGAGACGCCAACAG 927
Db 1031 GTAATGAAGAAAGAAACAGAGACGCCAACAG 1062

RESULT 4
LOCUS AX166534 1083 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 25 from Patent WO0138503.
ACCESSION AX166534
VERSION AX166534.1 GI:14546879
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clary,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0138503-A 25 31-MAY-2001;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source
1..1083
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

BASE COUNT	366 a	199 c	225 g	293 t	
ORIGIN					
Query Match	91.3%; Score 863; DB 6; Length 1083;				
Best Local Similarity	94.5%; Pred. No. 4.9e-197;				
Matches 927:	Conservative 0; Mismatches 0; Indels 54; Gaps 1;				
QY	1	ATGGAAAGTATGAAATAATTAGCTAGACCTGGAGAGGGCTTATGGGGTTGTATTCAA	60		
DB	1	ATGGAAAGTATGAAATAATTAGCTAGACCTGGAGAGGGCTTATGGGGTTGTATTCAA	60		
QY	61	TGCAGAAACAAACCTCTCGACAAAGTAGTAGCTGTTAAAAAATTTCTGGAATCTGAAGAT	120		
DB	61	TGCAGAAACAAACCTCTCGACAAAGTAGTAGCTGTTAAAAAATTTCTGGAATCTGAAGAT	120		
QY	121	GATCCTCTGTTTAAAGAAATAGCACTAAGAGAAATACGTAATGTTGAAGCAATTAACAAT	180		
DB	121	GATCCTCTGTTTAAAGAAATAGCACTAAGAGAAATACGTAATGTTGAAGCAATTAACAAT	180		
QY	181	CCAAATCTTCTGAACTCATCGAGGTGTTCCAGAGAAAGGAAATGCAATTTAGTTTTT	240		
DB	181	CCAAATCTTCTGAACTCATCGAGGTGTTCCAGAGAAAGGAAATGCAATTTAGTTTTT	240		
QY	241	GAATACCTGATCATACACATTTTAAATGAGCTGGAAGAAACCCAAATGAGTTGCTGAT	300		
DB	241	GAATACCTGATCATACACATTTTAAATGAGCTGGAAGAAACCCAAATGAGTTGCTGAT	300		
QY	301	GGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAATTTCTGTCTATATACAT	360		
DB	301	GGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAATTTCTGTCTATATACAT	360		
QY	361	AACGTATTCCACAGAGATATAAACCCTGGAATATCTATATACCTAAGCAAGAAATAATC	420		
DB	361	AACGTATTCCACAGAGATATAAACCCTGGAATATCTATATACCTAAGCAAGAAATAATC	420		
QY	421	AAGATTGTGACCTTCGGGTTTGCACAAAATCTCTGA	454		
DB	421	AAGATTGTGACCTTCGGGTTTGCACAAAATCTCTGAGTAATTCATATGTTGGAGAAATGAC	480		
QY	455	-----TTCCAGGAGATGCCTACACCGATTATGCTAGCT	486		
DB	481	TTACCTGATCTGATTGATGCAATTTGCGAGTTCCAGGAGATGCCTACACCGATTATGCTAGCT	540		
QY	487	ACGAGATGTTACCGAGCTCTGCACTTCTTGTGGGAGATACCTCAGTATGTTCTTCAGTC	546		
DB	541	ACGAGATGTTACCGAGCTCTGCACTTCTTGTGGGAGATACCTCAGTATGTTCTTCAGTC	600		
QY	547	GATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTGTGGCCT	606		
DB	601	GATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTGTGGCCT	660		
QY	607	GGAAATCAGATGTGGACCAACTTTATCTGATTAATCAGAACACTAGSAAATTAATCCCA	666		
DB	661	GGAAATCAGATGTGGACCAACTTTATCTGATTAATCAGAACACTAGSAAATTAATCCCA	720		
QY	667	AGACATCAATCAATCTTTAAAGTAACGGGTTTTCCATGSCATCAGTATACCTGAGCCA	726		
DB	721	AGACATCAATCAATCTTTAAAGTAACGGGTTTTCCATGSCATCAGTATACCTGAGCCA	780		
QY	727	GAAGACATGGAACCTCTTGAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACTTC	786		
DB	781	GAAGACATGGAACCTCTTGAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACTTC	840		
QY	787	ATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAG	846		
DB	841	ATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAG	900		
QY	847	AGCTCCTACTTTGATCTTTTCAAGAGGCCCAATTAAGAAAGAACACGTAATGAAGGA	906		
DB	901	AGCTCCTACTTTGATCTTTTCAAGAGGCCCAATTAAGAAAGAACACGTAATGAAGGA	960		
QY	907	AGAAACAGAGAGCGCCAACAG	927		

Db	961 AGAAACAGAGAGCCCAACAG 981				
RESULT 5					
AX107720					
LOCUS	AX107720	1041 bp	DNA	linear	PAT 30-APR-2001
DEFINITION	Sequence 9 from Patent WO0123579.				
ACCESSION	AX107720				
VERSION	AX107720.1	GI:13923201			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.				
AUTHORS	Human kinase proteins and polynucleotides encoding the same				
TITLE	Patent: WO 0123579-A 9 05-APR-2001;				
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
FEATURES	Lexicon Genetics Incorporated (US)				
source	Location/Qualifiers				
	1..1041				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	344 a	190 c	228 g	279 t	
ORIGIN					
Query Match	88.8%; Score 839; DB 6; Length 1041;				
Best Local Similarity	90.8%; Pred. No. 1.5e-181;				
Matches 945:	Conservative 0; Mismatches 0; Indels 96; Gaps 1;				
QY	1	ATGAAAAGTATCAAAAATTAGCTAAGACTCGAGAGGGCTTTATGGGGTTGTATTCAA	60		
DB	1	ATGAAAAGTATGAAAAATTAGCTAGACTGGAAGGGCTTTATGGGGTTGTATTCAA	60		
QY	61	TGCAGAAACAAACCTCTGGACAAAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120		
DB	61	TGCAGAAACAAACCTCTGGACAAAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120		
QY	121	GATCCTGCTGTTAAGAAATAGCACTAAGAGAAATACGTAATGTTGAAGCAATTAACAAT	180		
DB	121	GATCCTGCTGTTAAGAAATAGCACTAAGAGAAATACGTAATGTTGAAGCAATTAACAAT	180		
QY	181	CCAAATCTTGTGAACCTCATCGAGGTGTTCCAGAGAAAGGAAATGCAATTTAGTTTTT	240		
DB	181	CCAAATCTTGTGAACCTCATCGAGGTGTTCCAGAGAAAGGAAATGCAATTTAGTTTTT	240		
QY	241	GAATACCTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAAGTTGCTGAT	300		
DB	241	GAATACCTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAAGTTGCTGAT	300		
QY	301	GGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAATTTCTGTCTATATACAT	360		
DB	301	GGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAATTTCTGTCTATATACAT	360		
QY	361	AACGTATTCCACAGAGATATAAACCTGAAATATTTCTAATACTAAGCAAGGAATTAATC	420		
DB	361	AACGTATTCCACAGAGATATAAACCTGAAATATTTCTAATACTAAGCAAGGAATTAATC	420		
QY	421	AAGATTGTGACCTTCGGGTTTGCACAAAATCTGATTCCAGGAGATGCTACACCGATTAT	480		
DB	421	AAGATTGTGACCTTCGGGTTTGCACAAAATCTGATTCCAGGAGATGCTACACCGATTAT	480		
QY	481	GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTATGGTTCT	540		
DB	481	GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTATGGTTCT	540		
QY	541	TCAGTTCGATATATGGGCTATTGGTTGTTTTTTCGAGAGCTCCTGACAGGCCAGCCACTG	600		
DB	541	TCAGTTCGATATATGGGCTATTGGTTGTTTTTTCGAGAGCTCCTGACAGGCCAGCCACTG	600		
QY	601	TGGCCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAAATCAGAACACT	650		


```
Db 601 TGGCTGGAAATCAGATGTGGACCACTTTATCTGATAATCAGAACACTAGTAGACG 650
Qy 651 -----
Db 661 GGGTTTCGCCATGTTGACCAAGCTGGTCTCGAACTCTTGACGTCAAGTGTATCCACCTGCC 720
Qy 661 -----AGGAAATTAATCCCAAGACATCAATCAATCTT 684
Db 721 GTAGCCTCTCAAGTGTGGAAATTAACAGGAAATTAATCCCAAGACATCAATCAATCTT 780
Qy 685 AAAAGTAACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATCGGAACTCTT 744
Db 781 AAAAGTAACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATCGGAACTCTT 840
Qy 745 GAGGAAAGTCTCAGATGTTTCATCTCTGGCTCTGAACCTTCATGAAGGGGTGTCTGAAG 804
Db 841 GAGGAAAGTCTCAGATGTTTCATCTCTGGCTCTGAACCTTCATGAAGGGGTGTCTGAAG 900
Qy 805 ATGAATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAGAGCTCTCTACTTTGATTCT 864
Db 901 ATGAATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAGAGCTCTCTACTTTGATTCT 960
Qy 865 TTTCAGAGGCCCCAAATTAAGAAAGACAGATGATGAGGAAAGACAGAGAGCCCAA 924
Db 961 TTTCAGAGGCCCCAAATTAAGAAAGACAGATGATGAGGAAAGACAGAGAGCCCAA 1020
Qy 925 CAGGTACTTCGCTCAAAAGT 945
Db 1021 CAGGTACTTCGCTCAAAAGT 1041

RESULT 6
AX107714
LOCUS AX107714 1068 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123579.
ACCESSION AX107714
VERSION AX107714.1 G:13923198
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and
AUTHORS Sards,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 3 05-APR-2001;
LEXICON Genetics Incorporated (US)
FEATURES
source
location/Qualifiers
1..1058
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 344 a 197 c 240 g 287 t
ORIGIN
Query Match 87.1%; Score 823; DB 6; Length 1068;
Best Local Similarity 90.6%; Pred.No.7e-178;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

Qy 1 ATGGAAAGTATGAAATTAAGTAAAGTCTGAGAGGGTCTTATGGGTTGTATTCAA 60
Db 1 ATGGAAAGTATGAAATTAAGTAAAGTCTGAGAGGGTCTTATGGGTTGTATTCAA 60
Qy 61 TGCAGAAACAAACCTCTCGAAGTAGTGTGTTAAATAATTTGTGAAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTCGAAGTAGTGTGTTAAATAATTTGTGAAATCTGAAGAT 120
Qy 121 GATCCTGTTTGAAGAAATAGCACTAGAGAAATACGTATCTTCAAGCAATTAACAT 180
Db 121 GATCCTGTTTGAAGAAATAGCACTAGAGAAATACGTATCTTCAAGCAATTAACAT 180
```

```
Qy 181 CCAATCTTGTGAACCTCATCGAGCTGTTTCAGAGAGAAAAGGAAAATGCATTTAGTTTT 240
Db 181 CCAATCTTGTGAACCTCATCGAGCTGTTTCAGAGAGAAAAGGAAAATGCATTTAGTTTT 240
Qy 241 GAATCTGTGATCATACACTTTTAAATGAGTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATCTGTGATCATACACTTTTAAATGAGTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Qy 301 GGAGTGATCAAAAGCGTATTATGCGCAACACTTCAAGCTCTTAAATTTCTCATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGCGCAACACTTCAAGCTCTTAAATTTCTCATATACAT 360
Qy 361 AACTCTATTTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420
Db 361 AACTCTATTTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420
Qy 421 AAGATTTGTGATCTCGGGTTTGCACAAATTTCTGATTCAGAGAGATGCTACACCGATTAT 480
Db 421 AAGATTTGTGATCTCGGGTTTGCACAAATTTCTGATTCAGAGAGATGCTACACCGATTAT 480
Qy 481 GTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGAGATACTCAGTATGGTTCT 540
Db 481 GTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGAGATACTCAGTATGGTTCT 540
Qy 541 TCAGTCGATATATGGGCTATTGTTGTTGTTTTCGAGAGCTCTGACAGGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGTTGTTGTTTTCGAGAGCTCTGACAGGCCAGCCACTG 600
Qy 601 TGGCCTGGAAATACAGATGTCGACCACTTTATCTGATTAATCAGAACACTT----- 650
Db 601 TGGCCTGGAAATACAGATGTCGACCACTTTATCTGATTAATCAGAACACTT----- 650
Qy 651 -----
Db 661 GGGTTTCGCCATGTTGACCAAGCTGGTCTCGAACTCTTGACGTCAAGTGTATCCACCTGCC 720
Qy 661 -----AGGAAATTAATCCCAAGACATCAATCAATCTT 684
Db 721 GTAGCCTCTCAAGTGTGGAAATTAACAGGAAATTAATCCCAAGACATCAATCAATCTT 780
Qy 685 AAAAGTAACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTTT 744
Db 781 AAAAGTAACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTTT 840
Qy 745 GAGGAAAGTCTCAGATGTTTCATCTCTGGCTCTGAACCTTCATGAAGGGGTGTCTGAAG 804
Db 841 GAGGAAAGTCTCAGATGTTTCATCTCTGGCTCTGAACCTTCATGAAGGGGTGTCTGAAG 900
Qy 805 ATGAATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAGAGCTCTCTACTTTGATTCT 864
Db 901 ATGAATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAGAGCTCTCTACTTTGATTCT 960
Qy 865 TTTCAGAGGCCCCAAATTAAGAAAGACAGATGATGAGGAAAGACAGAGAGCCCAA 924
Db 961 TTTCAGAGGCCCCAAATTAAGAAAGACAGATGATGAGGAAAGACAGAGAGCCCAA 1020
Qy 925 CAGGT 929
Db 1021 CAGGT 1025

RESULT 7
AX286069
LOCUS AX286069 587 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 3 from Patent WO0179488.
ACCESSION AX286069
VERSION AX286069.1 GI:17045995
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```

AUTHORS      Kapeller-Libermann,R.
TITLE        14257, protein kinase molecules and uses therefor
JOURNAL      Patent: WO 0179488-A 3 25-OCT-2001;
              Millennium Pharmaceuticals, Inc. (US)
FEATURES     Location/Qualifiers
              source
              1..887
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              230 a 153 g 196 t
BASE COUNT   230 a 153 g 196 t
ORIGIN
1
230 a 153 g 196 t
Query Match      68.5%; Score 647; DB 6; Length 687;
Best Local Similarity 99.2%; Pred. No. 1.4e-137;
Matches 650; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGGAAGATGAAAAATTAGCTAAGCTGGAGAGGGCTTATGGGGTGTGATTCAAA 60
DB 1 ATGGAAGATGAAAAATTAGCTAAGCTGGAGAGGGCTTATGGGGTGTGATTCAAA 60
QY 61 TGCAGAACAAACCTCTGGACAAAGTAGTAGCTGTGTTAAAAAATTTGGGAATCTGAAGAT 120
DB 61 TGCAGAACAAACCTCTGGACAAAGTAGTAGCTGTGTTAAAAAATTTGGGAATCTGAAGAT 120
QY 121 GATCCTCTGTTAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTTAAACAT 180
DB 121 GATCCTCTGTTAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTTAAACAT 180
QY 181 CCAAACTCTGTGAACCTCATCGAGGTGTTCCAGGAGAAATAGCAATTTAGTTTTT 240
DB 181 CCAAACTCTGTGAACCTCATCGAGGTGTTCCAGGAGAAATAGCAATTTAGTTTTT 240
QY 241 GAATACCTGATCATACACTTTAAATGAGCTGGAAGAAACCCAAATGAGTTGCTGAT 300
DB 241 GAATACCTGATCATACACTTTAAATGAGCTGGAAGAAACCCAAATGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGCAACACTTCAGCTCTTAATTTCTGTCATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGCAACACTTCAGCTCTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420
QY 421 AAGATTTGTGACITCGGGTTTGCAAAATTTCTGATCCAGGAGATGCTACACCGATTAT 480
DB 421 AAGATTTGTGACITCGGGTTTGCAAAATTTCTGATCCAGGAGATGCTACACCGATTAT 480
QY 481 GTAGCTACGAGATGTTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTAGTATGTTCT 540
DB 481 GTAGCTACGAGATGTTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTAGTATGTTCT 540
QY 541 TCAGTGCATATATGGGTATTTGGTTGTTTTCAGAGCTCCTGACAGCCAGCCACTG 600
DB 541 TCAGTGCATATATGGGTATTTGGTTGTTTTCAGAGCTCCTGACAGCCAGCCACTG 600
QY 601 TGGCCTGGAAATCAGATGTGGACCACTTTATCTGATATCAGACACTAGGAA 655
DB 601 TGGCCTGGAAATCAGATGTGGACCACTTTATCTGATATCAGACACTAGGTA 655

RESULT 8
AX286067 LOCUS AX286067 382 bp DNA 1:near PAT 20-NOV-2001
DEFINITION Sequence 1 from Patent WO0179488.
ACCESSION AX286067
VERSION AX286067.1 GI:17045993
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
```

LOCUS AX056404 911 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 48 from Patent WO0073469.
ACCESSION AX056404
VERSION AX056404.1 GI:12229111
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
TITLE Protein Kinases
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source
1..911
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 293 a 182 c 178 g 258 t
ORIGIN
Query Match 68.0%; Score 643; DB 6; Length 911;
Best Local Similarity 100.0%; Pred. No. 1.le-136; Mismatches 0; Indels 0; Gaps 0;
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 285 AAATGGAGTTGCTGATGGAGTGCATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAA 344
Db 1 AAATGGAGTTGCTGATGGAGTGCATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAA 60
QY 345 TTTCGTGCATATACATAACTGTATTACAGAGATATAAACTTCAAAATATTCTAATTAAC 404
Db 61 TTTCGTGCATATACATAACTGTATTACAGAGATATAAACTTCAAAATATTCTAATTAAC 120
QY 495 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCACAGAGA 464
Db 121 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCACAGAGA 180
QY 465 TGCCTACACCGATTATGTAGCTACAGATGGTACCGAGCTCCTGAACCTTTCTGTGGGAGA 524
Db 181 TGCCTACACCGATTATGTAGCTACAGATGGTACCGAGCTCCTGAACCTTTCTGTGGGAGA 240
QY 525 TACTCAGTATGGTCTTCTCAGTCATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCT 584
Db 241 TACTCAGTATGGTCTTCTCAGTCATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCT 360
QY 585 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 644
Db 301 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 360
QY 645 AACCTTAGGAAATTAATCCCAAGACATCAATCAATCTTTAAAGTAAACGGGTTTTCCTCA 704
Db 361 AACCTTAGGAAATTAATCCCAAGACATCAATCAATCTTTAAAGTAAACGGGTTTTCCTCA 420
QY 705 TGGCATCAGTATACCTGAGCCAGAGACATGAAACTCTTTCAGGAAAGTTCTCAGATGT 764
Db 421 TGGCATCAGTATACCTGAGCCAGAGACATGAAACTCTTTCAGGAAAGTTCTCAGATGT 480
QY 765 TCATCTGTGGCTCTGAACCTTCATGAAGGGTGTCTGAAGATGATCCAGATGACAGATT 824
Db 481 TCATCTGTGGCTCTGAACCTTCATGAAGGGTGTCTGAAGATGATCCAGATGACAGATT 540
QY 825 AACCTGTTCCCAACTCCTGAGAGCTCTACTTTGATTCTTTTCAAGAGGCCCAAAATTA 884
Db 541 AACCTGTTCCCAACTCCTGAGAGCTCTACTTTGATTCTTTTCAAGAGGCCCAAAATTA 600
QY 885 AAGAAAGACAGTAAATGAAGGAAGAAACAGAGAGACGCCCAACAG 927
Db 601 AAGAAAGACAGTAAATGAAGGAAGAAACAGAGAGACGCCCAACAG 643

RESULT 10
AX056405

LOCUS AX056405 2615 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 49 from Patent WO0073469.
ACCESSION AX056405
VERSION AX056405.1 GI:12229112
KEYWORDS Murinae gen. sp.
SOURCE Murinae gen. sp.
ORGANISM Murinae gen. sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
REFERENCE 1
AUTHORS Plowman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
TITLE Protein Kinases
JOURNAL Patent: WO 0073469-A 49 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source
1..2615
/organism="Murinae gen. sp."
/mol_type="genomic DNA"
/db_xref="taxon:39108"
BASE COUNT 742 a 580 c 585 g 708 t
ORIGIN
Query Match 59.2%; Score 559.2; DB 6; Length 2615;
Best Local Similarity 81.8%; Pred. No. 1.5e-117; Mismatches 143; Indels 3; Gaps 1;
Matches 558; Conservative 0; Mismatches 143; Indels 3; Gaps 1;
QY 135 GAAATAGACCTAAGAGAAATACGTATGTTGAAGCAATTAAGACATCCAAATCTTGAA 194
Db 1 GAAATAGACCTCGGGAAATCCGTATGCTGAAG---TTGAACACCCAAACCTCGTGAA 57
QY 195 CCTCATCCAGGTGTTTCAGGAGAAAAGAAATCCATTTAGTTTTCGAATCTGTGATCA 254
Db 58 CTTATCCAGGTGTTTCAGGAGAAAAGAAAGATGCATCTAGTTTTGAGTACTGTGATCA 117
QY 255 TACACTTTTAAATCAGCTGGAAAGAAACCCAAATGGAGTTGCTGATGGAGTATCAAAAG 314
Db 118 CACACTGTTAAACAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGGAGTATTAAG 177
QY 315 CGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACATTAATCTGATTCACAG 374
Db 178 TGTGTTATGGCAACCCCTTCAAGCCCTTAACCTTCTGTCAAGCAATTTGATTCATCG 237
QY 375 AGATATAAAACCTGAAATATTTCTAATAACCTAAGCAAGGAATAATCAAGATTTCTGACTT 434
Db 238 GATGTTAAACCTGAAACATCTTAATAACCAAGAGGAGATGATAAGATTGTGACTT 297
QY 435 CGGGTTTCGCAAAATTCGTATTCAGAGATGCTTACACCGATTATGTAGTACGAGATG 494
Db 298 TGGATTTGCAGAAATTTCTAATTCAGGAGACGCTTACACAGACTATGTTGCCACAGGTG 357
QY 495 GTACCGAGCTCTGAACTTCTTGTGGAGATACCTCAGTATGTTCTTCAGTCCGATATG 554
Db 358 GTACCGAGCTCTGAACTTCTTGTGGAGACACCAAGTACGGTTCTCTGTAGACGTG 417
QY 555 GGTATTGTTGTTTGTGAGAGCTCTTGAAGCCAGCCACTGTGGCTGGAATAATC 614
Db 418 GGGCTGCGCTGTGTTTGTGAGAGCTCTTGAAGCCAGCCACTGTGGCTGGAATAATC 477
QY 615 AGATCTGCAACCACTTTATCTGATAATCAGAACACTAGGAAATTAATCCAGACATCA 674
Db 478 CAGCTGACAGCTTTTACCTGATCATCAGACCTTGGGGAAGCTGATTCAGACACCA 537
QY 675 ATCAATCTTTAAAGTAAACGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACAT 734
Db 538 GTCTATCTTTAGGAGTAAACAGTTTTCGCGGATCAGCATACCTGAAACAGAGGACAT 597
QY 735 GGAACACTCTTGAGGAAAGTTCTCAGATGTTTCATCTCTGTGCTCTGAACTTTCATGAAGGG 794
Db 598 GGAGACTCTTGAGGAAAGTTCTCAGATGTTTCATCTCTGTGCTCTGAACTTTCATGAAGGG 657
QY 795 GTCTCTGAGATGATCCAGATGACAGATTAACTGTTTCCCACTCTCGGAGAGCTCTTA 854
Db 658 ATGCTTGAAGATGATCTCTGATGAGGCTGACCTGTGCTGCTGCTGCTGCTGCTGCTGCT 717

```

QY 855 CTTTGATTCTTTCAAGAGCCCAATTAAGAGAAAGCAGTAATCGAGAGAGACAG 914
Db |||||
QY 718 CTTTGAGTCCTTTCAAGAGGATCAATGAAAGAGAAAGCCCGAGTCAGGGGAGAGCCG 777
QY 915 AAGACGGCAACACAGGTACTTCCGT 938
Db |||||
QY 778 AAGGCGCCAGCAGATCAACTGCT 801
Db |||||

RESULT 11
AX107724 1819 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 13 from Patent WO0123579.
ACCESSION AX107724
VERSION AX107724.1 GI:13923203
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and
Sands, A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 13 35-APR-2001;
FEATURES Lexicon Genetics Incorporated (US)
Source Location/Qualifiers
1. 1819
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 539 a 368 c 406 g 505 t : others
ORIGIN

Query Match 53.7%; Score 507; DB 6; Length 1819;
Best Local Similarity 69.8%; Pred. No. 1.4e-105;
Matches 929; Conservative 0; Mismatches 0; Indels 402; Gaps 2;

QY 1 ATGGAAGAGTATGAAAAATTAGCTAAGACTGGAGAGGGCTTTATGGGTTGTATCAAA 60
Db :38 ATGGAAGAGTATGAAAAATTAGCTAAGACTGGAGAGGGCTTTATGGGTTGTATCAAA :97
QY 61 TGCAGAAACAAACCTCTGGCAAGTAGTACCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 198 TGCAGAAACAAACCTCTGGCAAGTAGTACCTGTTAAAAAATTTGTGGAATCTGAAGAT 257
QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAAACAT 180
Db 258 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAAACAT 317
QY 181 CCAATCTGTGAACTCATCGAGTGTTTCAGGAGAAAGAAATGCAATTTAGTTTTT 240
Db 318 CCAATCTGTGAACTCATCGAGTGTTTCAGGAGAAAGAAATGCAATTTAGTTTTT 377
QY 241 GAATACCTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Db 378 GAATACCTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 437
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
Db 438 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 497
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAAGCAAGGAATAATC 420
Db 498 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAAGCAAGGAATAATC 557
QY 421 AAGATTGTGACTTCGGGTTTCACAAATTCGT----- 453
Db 558 AAGATTGTGACTTCGGGTTTCACAAATTCGTAGTTGGACTTCATCTTCTCTGGTGGC 617
QY 454 ----- 453

```

```

Db 618 TCCITGATTGGCTTAATAGTGTGACCTTCTGAATCTTTTTCGCAATTTCAGAGATTTT 677
QY 454 ----- 453
Db 678 CTCCTGCTTGGATTCATTCGTGACACAGTGTTCACACATGGGCCCCAGGCTCATCTCGA 737
QY 454 ----- 453
Db 738 ACTTCTGGCCTCAAGTGATCCTTCCACCTCGGCTCCCAAAGTGTCTGGATTGCAAGTGTG 797
QY 454 ----- 453
Db 798 AGCCACCGTGGCCAGCCAGATTTTTCACAAATTAATCTAGAGAGCTCACAAGATTGTTT 857
QY 454 -----ATTCCAGGAGATGCTTACACC 474
Db 858 TTAGTGGGACACAATTTTCGACAAATTTCTTGAGAACGCAATTCAGGAGATGCTTACACC 917
QY 475 GATTATGATGCTACGAGATGGTACCGAGCTCTCGAACTTCTTGTGGAGATATCTAGTAT 534
Db 918 GATTATGATGCTACGAGATGGTACCGAGCTCTCGAACTTCTTGTGGAGATATCTAGTAT 977
QY 535 GGTCTTTCAGTCGATATATGGGCTATTGGTGTGTTTTCGACAGCTCCTGCAGAGCCAG 594
Db 978 GGTCTTTCAGTCGATATATGGGCTATTGGTGTGTTTTCGACAGCTCCTGCAGAGCCAG 1037
QY 595 CCACCTGTGGCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACT--- 650
Db :039 CCACCTGTGGCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA 1097
QY 651 ----- 650
Db :038 GAGACGGGGTTTCGCCATGTTGACCCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCA 1157
QY 651 -----AGGAAATTAATCCCAAGACATCAATCA 678
Db 1158 CCGCGCTAGCCTCTCAAAGTGTGGAATTAACAGGAAATTAATCCCAAGACATCAATCA 1217
QY 679 ATCTTTAAAGTAACGGGTTTTCCTATGGCATCATGATAGTCTGAGCCAGAGACATGGAA 738
Db 12-8 ATCTTTAAAGTAACGGGTTTTCCTATGGCATCATGATAGTCTGAGCCAGAGACATGGAA 1277
QY 739 ACTCTTGAGGAAAGTCTCAGATGTTTCATCTCGCTGCTGAACTTCTGAACTTCTGAGGGGTGT 798
Db 1278 ACTCTTGAGGAAAGTCTCAGATGTTTCATCTCGCTGCTGAACTTCTGAACTTCTGAGGGGTGT 1337
QY 799 CTGAAGATGAATCCAGATGACAGATTAACCTGTTTCCCAACTCTCTGGAGAGCTCTACTTT 858
Db 1338 CTGAAGATGAATCCAGATGACAGATTAACCTGTTTCCCAACTCTCTGGAGAGCTCTACTTT 1397
QY 859 GATCTTTTCAGAGGGCCCAATTAAGAGAAAGCAGTAATGAAGGAGAGAAACAGAGA 918
Db 1398 GATCTTTTCAGAGGGCCCAATTAAGAGAAAGCAGTAATGAAGGAGAGAAACAGAGA 1457
QY 919 CGCCCAACAGGT 929
Db :458 CGCCCAACAGGT 1468

RESULT 12
AX107712 561 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 1 from Patent WO0123579.
DEFINITION AX107712
ACCESSION AX107712
VERSION AX107712.1 GI:13923197
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and
Sands, A.T.

```

TITLE		Human kinase proteins and polynucleotides encoding the same	
JOURNAL		Patent: WO 0123579-A 1 05-APR-2001;	
Lexicon Genetics Incorporated (US)		Lexicon Genetics Incorporated (US)	
FEATURES	source	Location/Qualifiers	
	1..561	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
BASE COUNT		189 a 187 c 112 g 173 t	
ORIGIN			
Query Match		48.1%; Score 454.4; DB 6; Length 561;	
Best Local Similarity		99.8%; Pred. No. 1.7e-93;	
Matches 455; Conservative		0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	ATGGAAGATGTAAGAAATTTAGCTTAAGACTGGAGAGGGCTCTTATCGGGTTGTATTCAAA	60
	1	ATGGAAGATGTAAGAAATTTAGCTTAAGACTGGAGAGGGCTCTTATCGGGTTGTATTCAAA	60
DB	61	TGCAGAAACAAACCTCTGGACAAAGTAGTGGTAAAGAAATTTTGTGGAATCTGAAGAT	120
	61	TGCAGAAACAAACCTCTGGACAAAGTAGTGGTAAAGAAATTTTGTGGAATCTGAAGAT	120
QY	121	GATCCTCTTGTGAAGAAATAGCACAATAGCAATAGCAATAGCAATAGCAATAGCAAT	180
	121	GATCCTCTTGTGAAGAAATAGCACAATAGCAATAGCAATAGCAATAGCAATAGCAAT	180
QY	181	CCAAATCTTGTGAACCTCATCGAGTGTTCAGAGAAAGGAAATGCAATTTAGTTTT	240
	181	CCAAATCTTGTGAACCTCATCGAGTGTTCAGAGAAAGGAAATGCAATTTAGTTTT	240
QY	241	GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGTGTGAT	300
	241	GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGTGTGAT	300
QY	301	GGAGTGATCAAAAGCGTATTATGGCAAAACCTTCAAGCTCTTAATTTCTGTCTATACAT	360
	301	GGAGTGATCAAAAGCGTATTATGGCAAAACCTTCAAGCTCTTAATTTCTGTCTATACAT	360
QY	361	AACTGTATTACAGAGATATAAACCCTGAAATATTTCTAATTAACCAAGCAAGAAATATC	420
	361	AACTGTATTACAGAGATATAAACCCTGAAATATTTCTAATTAACCAAGCAAGAAATATC	420
QY	421	AAGATTGTGACTTCGGGTTTGCAAAATTCGATT	456
	421	AAGATTGTGACTTCGGGTTTGCAAAATTCGATT	456
RESULT 14			
AX107718		1612 bp DNA linear PAT 06-JUN-2002	
LOCUS		Sequence 146 from Patent WO0218424.	
DEFINITION		AX107718	
ACCESSION		AX107718	
VERSION		AX107718.1 GI:21336301	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1	
AUTHORS		Tang, Y. T., Asundi, V., Zhou, P., Xue, A. J., Ren, F., Zhang, J., Wang, J. R., Zhao, Q. A., Wang, D., Liu, C., Drmanac, R. T. and Wehrman, T.	
TITLE		Nucleic acids and polypeptides	
JOURNAL		Patent: WO 0218424-A 146 07-MAR-2002;	
FEATURES		HySEQ, INC. (US)	
source		Location/Qualifiers	
		1..1612	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		463..1539	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD13419.1"	
		/db_xref="GI:21336302"	
		/translation="MMEKYKIKIGEGSVGVFKCRNBDTQIVAKKFLSEDDPV	
		LKIALREIRMLKQLKHPNLVNLLEFRRLHLVFEYCDHTVLHLDRLYGRGVPEH	
		LVKSIITWQLQAVNFCRHNCHIRDKVKNPILITKHSVILKLCDFGFLRLTGPDSYIT	
		DYATRWYSPPELLVGTQYGPVVDVWAGTGVFAELLSGVLPWPGSDVDQLYLRKT	
		LGDLIPIHQOVSFNQYFSGVKIPDPDMEPLELKFPI SYPALGLKGLMDPFR	
		LTCEOLLHPYFENIREIELDAKHNKPTKTLKSRKHHCFTETSKLYLPQLTGSS	
		ILPALDNKKYYCDTKLNLNRPNI"	
BASE COUNT		476 a 380 c 361 g 395 t	
ORIGIN			

TITLE		Human kinase proteins and polynucleotides encoding the same	
JOURNAL		Patent: WO 0123579-A 1 05-APR-2001;	
Lexicon Genetics Incorporated (US)		Lexicon Genetics Incorporated (US)	
FEATURES	source	Location/Qualifiers	
	1..561	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
BASE COUNT		189 a 187 c 112 g 173 t	
ORIGIN			
Query Match		48.1%; Score 454.4; DB 6; Length 561;	
Best Local Similarity		99.8%; Pred. No. 1.7e-93;	
Matches 455; Conservative		0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	ATGGAAGATGTAAGAAATTTAGCTTAAGACTGGAGAGGGCTCTTATCGGGTTGTATTCAAA	60
	1	ATGGAAGATGTAAGAAATTTAGCTTAAGACTGGAGAGGGCTCTTATCGGGTTGTATTCAAA	60
DB	61	TGCAGAAACAAACCTCTGGACAAAGTAGTGGTAAAGAAATTTTGTGGAATCTGAAGAT	120
	61	TGCAGAAACAAACCTCTGGACAAAGTAGTGGTAAAGAAATTTTGTGGAATCTGAAGAT	120
QY	121	GATCCTCTTGTGAAGAAATAGCACAATAGCAATAGCAATAGCAATAGCAATAGCAAT	180
	121	GATCCTCTTGTGAAGAAATAGCACAATAGCAATAGCAATAGCAATAGCAATAGCAAT	180
QY	181	CCAAATCTTGTGAACCTCATCGAGTGTTCAGAGAAAGGAAATGCAATTTAGTTTT	240
	181	CCAAATCTTGTGAACCTCATCGAGTGTTCAGAGAAAGGAAATGCAATTTAGTTTT	240
QY	241	GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGTGTGAT	300
	241	GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGTGTGAT	300
QY	301	GGAGTGATCAAAAGCGTATTATGGCAAAACCTTCAAGCTCTTAATTTCTGTCTATACAT	360
	301	GGAGTGATCAAAAGCGTATTATGGCAAAACCTTCAAGCTCTTAATTTCTGTCTATACAT	360
QY	361	AACTGTATTACAGAGATATAAACCCTGAAATATTTCTAATTAACCAAGCAAGAAATATC	420
	361	AACTGTATTACAGAGATATAAACCCTGAAATATTTCTAATTAACCAAGCAAGAAATATC	420
QY	421	AAGATTGTGACTTCGGGTTTGCAAAATTCGATT	456
	421	AAGATTGTGACTTCGGGTTTGCAAAATTCGATT	456
RESULT 13			
AX107718		594 bp DNA linear PAT 30-APR-2002	
LOCUS		Sequence 7 from Patent WO0123579.	
DEFINITION		AX107718	
ACCESSION		AX107718	
VERSION		AX107718.1 GI:13923200	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1	
AUTHORS		Donoho, G., Turner, C. A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A. T.	
TITLE		Human kinase proteins and polynucleotides encoding the same	
JOURNAL		Patent: WO 0123579-A 7 05-APR-2001;	
Lexicon Genetics Incorporated (US)		Lexicon Genetics Incorporated (US)	
FEATURES	source	Location/Qualifiers	
	1..594	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
BASE COUNT		203 a 94 c 115 g 182 t	
ORIGIN			

Query Match	44.0%	Score 415.4	DB 6	Length 1612
Best Local Similarity	56.7%	Pred. NC. 1.3e-84		
Matches 609; Conservative	0	Mismatches 301	Indels	Gaps 1
QY	1	ATGGAAGATGATCAAAATTTAGCTAAAGCTGGAGAGGGTCTTATGGGTGTGATTCAAA	60	
DB	466	ATGGAGAGTATGAAAAATTTGGAAATTTGGAGAGGATCCTATGGAGTGTGTTTCAAA	525	
QY	61	TGCAGAAACAAACCTCTGACCAAGTAGTAGCTGTCTAAAAAATTTGTGGAATCTGAAGAT	120	
DB	526	TGTAGAAACAGGGACACGGGTCAAGTTGTGGCCATCAAGAAGTTCTTGGAAATCAGAAGAT	585	
QY	121	GATCCGTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATCTGTGAGCAATTAACAAT	180	
DB	586	GACCCGTGCTAAGAAATTTGCCCTTCGGGAAATCGGAATGCTCAAGCAACTCAAGCAT	645	
QY	181	CCAAATCTTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATATGATTTAGTITTT	240	
DB	646	CCCAACTTTGTTAACTCTCTGGAAGCTTCAGGAGGAAACGGAGGCTTCACTTGGTGT	705	
QY	241	GATACCTGTGATACATACCTTTTAAATGAGCTGGAAAGAAACCCAAATCGAGTGTGCTGAT	300	
DB	706	GAAATTTGACCAACACAGTTCTCCATGAGTTGGACAGATACCAAGAGGGGTACAGAA	765	
QY	301	GGAGTGATCAAAAGCCTATTATGGCAAACTTCAAGCTCTTAATTTCTGTCTATATACAT	360	
DB	766	CATCTCGTGAAGAGCATAACTTGGACACACTGCAAGCTGTAATTTTGCCATAAACAC	825	
QY	361	AACTGTATTTACAGAGATATAAAACCTGAAATATTTCTAATACATAAGCAAGAAATATC	420	
DB	826	AATTTGATACATAGACAGCTGAAGCCAGAAATATCTCTCATCAGAAACATTCCTGTGAT	885	
QY	421	AAGATTTGTGACCTTCGGGTTTGACAAATTTCTGATT--CCAGGAGATGCTACACCGAT	477	
DB	886	AAGCTTTGTGACCTTGGATTTGCTCGCTTTGACTGGACCGAGTGACTACTATACAGAC	945	
QY	478	TATGTAGCTACGAGATGGTACCAGCTCTCGAACTCTTCTGTGGAGATACCTAGTATGTT	537	
DB	946	TAGTGGCTACCAAGTGGTACCGCTCCCTGAGCTGCTGGTGGGGACACGGCAGTACGGC	1005	
QY	538	TCTTCAGTCGATATATGGGCTATTTGGTGTGTTTTTCAGAGCTCTCTGACAGCCAGCCA	597	
DB	1006	CCCCCGGTGATGTTTGGGCAATTTGGCTGTCTTTGCTGAGTGTCTGTGAGGATGCTCT	1065	
QY	598	CTGTGCTCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAA	657	
DB	1066	CTGTGGCCAGAAAAATCGAGTGGATCGACTGTATCTGATTAGGAAGACCTTGGGGAT	1125	
QY	658	TTAATCCCAAGACATCAATCAATCTTTAAAGTAAACGGGTTTTTCATGSCATCAGTATA	717	
DB	1126	CTCATTTCTTAGGCACAGCAAGTGTTTAGCACGAATCAGTACTTCAGTGGAGTGAATAAT	1185	
QY	718	CCTGAGCCAGAACATGGAATCTTTGAGGAAAAAGTTCTCAGATGTTTCATCTCTGGCT	777	
DB	1186	CCAGACCTCTGAAGATATGGAACCACTTGAATTAATTTCCAAACATCTCTTATCCTGCC	1245	
QY	778	CTGAATCTCATGAAGGGTGTCTGAAGATGAATTCAGATGACAGATTAACCTGTTCCCAA	837	
DB	1246	CTGGGGCTCCTAAAGGGCTGTCTCCATGATGGACCTCATCTGAAAGGGCTGACATGTGAACAG	1305	
QY	838	CTCCTGGAGAGCTCCTACTTTGATTTCTTTTCAAGAGGCCCAATTAAGAAAAAGACAGT	897	
DB	1306	CTGTTGATCACCCTATATTTTGAACATCAGAGAAATAGAGGATTTGGCAAAAGAACAC	1365	
QY	898	AATGAAGGAAGAA	910	
DB	1366	AACAAAACCAACA	1378	

RESULT 15
HSSTHPKB
LOCUS

DEFINITION H. sapiens mRNA KXIALRE for serine/threonine protein kinase.
ACCESSION X663358
VERSION X663358.1 GI:36614
KEYWORDS serine/threonine protein kinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1363)
AUTHORS MeyerSON,M., Enders,G.H., Wu,C.L., Su,L.K., Gorka,C., Nelson,C.,
Harlow,E. and Tsai,L.H.
TITLE A family of human cdc2-related protein kinases
JOURNAL EMBO J. 11 (8), 2909-2917 (1992)
MEDLINE 92347325
PUBMED 1639063
REFERENCE 2 (bases 1 to 1363)
AUTHORS MeyerSON,M.L.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1992) M.L. MeyerSON, Massachusetts General
Hospital, Cancer Center, Bldg 149, 13th Street, Charlestown MA
02129, USA
FEATURES Location/Qualifiers
source 1..1363
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
cell_line="HeLa"
cell_type="HeLa"
clone_lib="HeLa"
214..1290
codon_start=1
product="serine/threonine protein kinase"
protein_id="CAA47002.1"
db_xref="GI:36615"
db_xref="SWISS-PROT:Q00532"
translation="MMEKYEKGKIGEGSYGVVFKCRNDRDGTQIVAIKKFLESDPVP
IKKAIWTLQKLVKRLKQNLVNLVFRKKRRLHLVFEYCDHTVLHLEDRYQGVPEH
LVKSIWOTLQVNFCHGKTHRDVKKRNLITKHSVILKDFDFAILLDTGQVSDYIT
DYVATWYRSPELLVGGDTGPPVDVMAICVFAELLSGVPLWPGKSDVDQLYLIRKT
LGLDI PRHQQVSTNVOYSGVKIPDPEDMEPLKFPNI:SYPALGILLKGLHMDPTER
LTCOOLLHHYPENIREIEDLAKEHDKPTRKTLRKSRKHCFETETSKLQVLPQLTGS
IIPALDNKKYCYCTKKLVYRFPNI"
join(214..569,X66359.1:1..54,670..1290)
notes="KXIALRE coding sequence with 54 bp insertion"
pseudo
/codon_start=1
/label=s:hpkin_CDS
product="serine/threonine protein kinase"
db_xref="PID:e49306"
BASE COUNT 412 a 308 c 309 g 334 t
ORIGIN
Query Match 43.8%; Score 413.8; DB 9; Length 1363;
Best Local Similarity 66.6%; Pred. No. 3e-84;
Matches 608; Conservative 0; Mismatches 302; Indels 3; Gaps 1;
QY 1 ATCGAAAGTATGAAAAATTAGCTAGACTGGAGAGGGTCATTATCGGGTTGATTCAAA 60
Db 217 ATGGAGAAGTATGAAAAAATTGGGAAAATGGGAGAAGGATCCTATGAGGTGTTTTCAAA 276
QY 61 TGCAGAAAACAAACCTCTGCAACAAGTAGTAGCTGTTTAAAAAAATTTCTGGAATCTGAAGAT 120
Db 277 TGTAGAAAACAGGACACGGGTGAGATTGTGGCCATCAAGAGTTTCTGNAATCAGAAGAT 336
QY 121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180
Db 337 GACCCGTGTCATTAAGAAAAATTCGCCCTTCGGGAAATCCGAATGCTCAAGCAACTCAAGCAT 396
QY 181 CCAAACTCTGTGAACCTCATCGAGGTGTTTCAGAGAAAAGGAAAATGCATTTAGTTTTT 240
Db 397 CCCAACCTTGTTAACCTCTCGGAAGTCTTCAGAGAGAAACGGAGGCTTCACCTGGGTGTTT 456
QY 241 GAATACTGTGATCATACACTTTTAAATAGAGCTGGAAGAGAAACCAATGGAGTTGCTGAT 300

457	db	GAATATTGTGACCAACAGATTCTCCATGAGTTGGACAGATACCAAGAGGGGTACCGAA	516
301	Qy	GGAGTGAACAAAGCGTTATTGCAAAACACTTCAAGCTCTTAATTTCTGTCTATATACAT	360
517	Db	CATCTCGTGAAGACATAAAGTGTGGACACACTGCAAGCTGTAAATTTTGGCATAAACAC	576
361	Qy	AAGCTGATTTCACAGAGATATAAAACCTGAAATTAATCTTAATTAACTAAGCAAGGAATATC	420
577	Db	AATTGCAATACATAGACAGCTGAAGCAGAAATATCCTCATCACGAAACATTTCCGTGATT	636
421	Qy	AAGATTTGTGACTTCGGGTITGCACAAATTTCTGATT---CCAGGAGATGCGCTACACCGAT	477
637	Db	AAGCTTTGTGACTTTGGATTGTCTCGGCTTTTGACTGGACCGAGTGACTACTATACAGAC	696
478	Qy	TATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACTCAGTNTGGT	537
697	Db	TACGTGGGCTACAGGTGGTACCGCTCCCGCTGAGCTGCTGTGGGGGACACGAGTAGCGC	756
538	Qy	TCTTCAGTTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCTGCACAGGCGACGCCA	597
757	Db	CCCCGGTGGATGTTTGGCAATTGGCTGTGCTCTTGTGAGCTCTGTGAGGAGTGCCT	816
598	Qy	CTGTGGCTGAAATACAGATGTGGACCACTTATCTGATTAATCAGAACACTAGGAAA	657
817	Db	CTGTGGCCAGAAATCGGATGTGGATGAGCTGTCTGATTTAGGAAGACCTTGGGGGAT	876
658	Qy	TTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTCCATGGCATCAGTATA	717
877	Db	CTCATTTCTAGGCCACGCAAGTGTTTAGCACCAATCAGTACTTCAGTGGAGTGAATAAT	936
718	Qy	CCTGAGCCAGAGACATGGAACTCTTGAGAAAGCTTCTCAGATGTTCACTCTGTGGCT	777
937	Db	CCAGACCTTGAGATATGGAACCACTTGAATTAATAATTCCAAAACATCTCTTATCTGCC	996
778	Qy	CTGAACCTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAA	837
997	Db	CTGGGGCTCTAAAGGGGTGTCTCCACATGGACCTTACTGAAAGGCTGCATGTGAACAG	1056
838	Qy	CTCTTGGAGGCTCTACTTTTGATTCTTTTCAAGAGGCCCAATTAAGAAGAAAGACGCT	897
1057	Db	CTGTGTGATCCCTATTTTGAANAACATCAGAGAAATAGAGATTTCGCAANAAGACAC	1116
898	Qy	AATGAAGGAAGAA	910
1117	Db	GACAAACCAACAA	1129

Search completed: November 15, 2003, 19:04:23
Job time : 3707.44 secs

1	945	100.0	945	22	AAD03817	Human: kinase CDNA
2	929	98.3	972	22	AAD03814	Human kinase CDNA
3	927	98.1	1790	24	AAD30557	Human kinase CDNA
4	868.8	91.9	1678	25	ABX08936	Human kinase polyA
5	863	91.3	1083	22	AA506725	CDNA encoding human
6	839	88.8	1041	22	AAAD03816	Polynucleotide seq
7	823	87.1	1068	22	AAD03816	Human kinase CDNA
8	645.4	68.3	982	24	AAI64248	Human kinase CDNA
9						Human kinase 14257

PT compounds useful for treating mental, biological or medical diseases -

PS Disclosure; Page 29-30; 38pp; English.

XX The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological, or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match 98.3%; Score 929; DB 22; Length 972;

Best Local Similarity 100.0%; Pred. No. 8.5e-241;

Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAGTATCAAAAATTAGCTAAGACTGAGAGAGGGTCTTATGGGGTGTATTCAAA 60
DB 1 ATGGAAGAGTATCAAAAATTAGCTAAGACTGAGAGAGGGTCTTATGGGGTGTATTCAAA 60
QY 61 TCGAAGAACAAACCTCTGGCAAGTAGTACTGTATTAATAATTTTGGGAATCTGAAGAT 120
DB 61 TCGAAGAACAAACCTCTGGCAAGTAGTACTGTATTAATAATTTTGGGAATCTGAAGAT 120
QY 121 GATCTCTGTTGTAAGAAATACACTAGAGAAATACGTATGTCAGCAATTAACACAT 180
DB 121 GATCTCTGTTGTAAGAAATACACTAGAGAAATACGTATGTCAGCAATTAACACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTAGTTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTAGTTTT 240
QY 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300
DB 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
DB 301 GGAGTGATCAAAAGCGGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
QY 361 AACTGTATTCACAGAGATATAAAGCTGAATATTTCTAATACTAAGCAAGGAATATTC 420
DB 361 AACTGTATTCACAGAGATATAAAGCTGAATATTTCTAATACTAAGCAAGGAATATTC 420
QY 421 AAGATTGTGACTTCGGGTTTCACAAATTCGTGATTCAGAGAGATGCCTACACCGATTAT 480
DB 421 AAGATTGTGACTTCGGGTTTCACAAATTCGTGATTCAGAGAGATGCCTACACCGATTAT 480
QY 481 GTAGCTACAGAGATGTGATCCGAGCTCTGAACCTCTTGTGGGAGATACATCATGGTTCT 540
DB 481 GTAGCTACAGAGATGTGATCCGAGCTCTGAACCTCTTGTGGGAGATACATCATGGTTCT 540
QY 541 TCAGTCGATATATGGCTATTGGTTGTGTTTTCAGAGAGCTCTTCGACAGGCCACCGACTG 600
DB 541 TCAGTCGATATATGGCTATTGGTTGTGTTTTCAGAGAGCTCTTCGACAGGCCACCGACTG 600

QY 601 TGGCTCTGAAAATCAGATGTGACCAACTTTTCTGATTAATCAGAACACTAGGAAAATTA 660
DB 601 TGGCTCTGAAAATCAGATGTGACCAACTTTTCTGATTAATCAGAACACTAGGAAAATTA 660
QY 661 ATCCCAAGACATCAATCAATCTTTAAAGAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
DB 661 ATCCCAAGACATCAATCAATCTTTAAAGAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
QY 721 GAGCAGAGAAGATCGGAAACTCTTCAGGAAAAGTTCTCAGATGTTTCACTCTGGCTCTG 780
DB 721 GAGCAGAGAAGATCGGAAACTCTTCAGGAAAAGTTCTCAGATGTTTCACTCTGGCTCTG 780
QY 781 AACTTCATGAAGGGGTGCTGAAGATGAATCCAGATCAGAGATTAACCTGTTCCCAACTC 840
DB 781 AACTTCATGAAGGGGTGCTGAAGATGAATCCAGATCAGAGATTAACCTGTTCCCAACTC 840
QY 841 CTGGAGAGCTCTACTTTGATTTCTTTTCAAGAGGCCCAAAATTAAGAAAAGCAGCTAAT 900
DB 841 CTGGAGAGCTCTACTTTGATTTCTTTTCAAGAGGCCCAAAATTAAGAAAAGCAGCTAAT 900
QY 901 GAGGAGAGAAACAGAGACGCGCAACAGGT 929
DB 901 GAGGAGAGAAACAGAGACGCGCAACAGGT 929

RESULT 3
AAD30557
ID AAD30557 standard; cDNA; 1790 BP.
XX AAD30557;
XX 21-MAY-2002 (first entry)
XX Human kinase polypeptide (PKIN-10) cDNA.
DE Human; kinase polypeptide; PKIN-10; gene therapy; Addison's disease;
XX leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
KW cirrhosis; hepatitis; peoriiasis; Cushing's syndrome; cytostatic; cancer;
KW cholestasis; cardiact; cardiovascular disorder; Niemann-Pick's disease;
KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
KW hypotensive; anti-HIV; enzyme; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
CDS 239..1267
FT /*tag= a
FT /product= "Human PKIN-10"
XX W0200208399-A2.
XX 31-JAN-2002.
XX 20-JUL-2001; 2001WO-US21092.
XX 21-JUL-2000; 2000US-220038P.
PR 28-JUL-2000; 2000US-222112P.
PR 04-AUG-2000; 2000US-222831P.
PR 11-AUG-2000; 2000US-224729P.
XX (INCY-) INCYTE GENOMICS INC.
PA (THOR-) THORNTON M.
XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Wallia NK;
PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P;
PI Ding L, Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM;
PI Greenwald SR, Tang YT, Xu Y, Walsh NT, Gietzen KJ, Yang J;
PI Hillman JL;

XX WPI: 2002-206083/26.
 DR P-PSDB; AAE19152.

XX New human kinase polypeptide, useful in diagnosis, prevention and
 PT treatment of cancer, immune disorder, growth and developmental
 PT disorder, cardiovascular disorder and lipid disorder -

XX Claim 5; Page 183; 196pp; English.

XX The present invention relates to an isolated human kinase polypeptide
 CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
 CC useful for diagnosing, treating and preventing cancer (e.g., leukaemia,
 CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
 CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
 CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
 CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
 CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
 CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
 CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
 CC drug screening techniques and to analyse the proteome of a tissue or cell
 CC type. PKIN is useful for creating knockin humanised animals or transgenic
 CC animals to model human diseases, in somatic or germline gene therapy, to
 CC generate a transcript image of a tissue or cell type, for detecting
 CC differences in the chromosomal location due to translocation, inversion,
 CC etc., among normal, carrier or affected individuals, and as hybridisation
 CC probes for mapping naturally occurring genomic sequences. PKIN is useful
 CC in southern or northern analysis, dot blot or other membrane-based
 CC technologies, in PCR technologies, in dipstick, pin, multiformat enzyme
 CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
 CC fluids or tissues from patients to detect altered PKIN expression. The
 CC present sequence is human PKIN-10 cDNA.

XX Sequence: 1790 BP; 581 A; 328 C; 366 G; 515 T; 0 other;

Query Match 98.1%; Score 927; DB 24; Length 1790;
 Best Local Similarity 100.0%; Pred. No. 3.7e-240;
 Matches 927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAGTATGAAATTAAGCTTAAGCTTGAAGAGGCTTATGGGTGTTGATCAAA 60
 Db 239 ATGAAAGTATGAAATTAAGCTTAAGCTTGAAGAGGCTTATGGGTGTTGATCAAA 298
 Qy 61 TCGAGAACCAAACTCTGACRAGTAGTACCTGTTAAATTTTCGGATTGAGAT 120
 Db 299 TCGAGAACCAAACTCTGACRAGTAGTACCTGTTAAATTTTCGGATTGAGAT 358
 Qy 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATAGTATGTTGAGCAATTAACAAT 180
 Db 359 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATAGTATGTTGAGCAATTAACAAT 418
 Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTGAGGAAATAGGAAATGCAATTTAGTTTT 240
 Db 419 CCAATCTTGTGAACCTCATCGAGGTGTTGAGGAAATAGGAAATGCAATTTAGTTTT 478
 Qy 241 GAATACGTGATCATACACTTTAAATGAGCTGAAAGAACCCAAATGAGTGGTGAT 300
 Db 479 GAATACGTGATCATACACTTTAAATGAGCTGAAAGAACCCAAATGAGTGGTGAT 538
 Qy 301 GGAGTGATCAAAAGCGTATTATGSCAAACACTTCAAGCTCTTAATTTCTGTCATACAT 360
 Db 539 GGAGTGATCAAAAGCGTATTATGSCAAACACTTCAAGCTCTTAATTTCTGTCATACAT 598
 Qy 361 AACTGTATTCAGAGATATAAACTCGAATATTTTAATTAAGCAAGAAATATC 420
 Db 599 AACTGTATTCAGAGATATAAACTCGAATATTTTAATTAAGCAAGAAATATC 658
 Qy 421 AAGATTTGTGACTTCGGGTTTGCAAAATCTCGATTCAGAGATGCTACACCGATTAT 480
 Db 659 AAGATTTGTGACTTCGGGTTTGCAAAATCTCGATTCAGAGATGCTACACCGATTAT 718
 Qy 481 GTAGTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTAGTATGGTCT 540

Db 719 GTAGCTACGAGATGGTACCAGAGCTCCTGAACTCTTGTGCGAGATACTCAGTATGGTTCT 778
 Qy 541 TCACTCGATATATGGGCTATTTGGTGTGTTTTCAGAGAGCTCTCGACAGGCGCAGCACTG 600
 Db 779 TCACTCGATATATGGGCTATTTGGTGTGTTTTCAGAGAGCTCTCGACAGGCGCAGCACTG 838
 Qy 601 TGGCTCGAAATACAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAATTA 660
 Db 839 TGGCTCGAAATACAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAATTA 898
 Qy 661 ATCCCAAGACATCAATCAATCTTTTAAAGTAAGCGGTTTTTCCATCGCATACATACCT 720
 Db 899 ATCCCAAGACATCAATCAATCTTTTAAAGTAAGCGGTTTTTCCATCGCATACATACCT 958
 Qy 721 GAGCCAGAACATCGGAACTCTTGAGGAAAGCTTCTCAGATGTTTCTGTTGGCTCTG 780
 Db 959 GAGCCAGAACATCGGAACTCTTGAGGAAAGCTTCTCAGATGTTTCTGTTGGCTCTG 1018
 Qy 781 AACTTCATGAAGGCGTCTCTGAAGATGAATCCAGATCAGACAGATTAACCTGTTCCCAACTC 840
 Db 1019 AACTTCATGAAGGCGTCTCTGAAGATGAATCCAGATCAGACAGATTAACCTGTTCCCAACTC 1078
 Qy 841 CTGGAGAGCTCTTACTTTTGATTTCTTTCAGAGGCCCAATTAAGAAAGAACGACCTAT 900
 Db 1079 CTGGAGAGCTCTTACTTTTGATTTCTTTCAGAGGCCCAATTAAGAAAGAACGACCTAT 1138
 Qy 901 GAAGGAAGAAACAGAGACGCCCAACAG 927
 Db 1139 GAAGGAAGAAACAGAGACGCCCAACAG 1165

RESULT 4
 ABX08936
 CD ABX08936 standard; cDNA; 1678 BP.
 XX AC ABX08936;
 XX CT 21-JAN-2003 (first entry)
 XX CD cDNA encoding human D1HP protein #6.
 DE Human; ss; gene; diagnostic and therapeutic; D1HP; cancer;
 KW arteriosclerosis atherosclerosis; psoriasis; primary thrombocytopenia;
 KW autoimmune; inflammatory; anaemia; asthma; autoimmune thyroiditis;
 KW Crohn's disease; diabetes mellitus; glomerulonephritis; gout; stroke;
 KW multiple sclerosis; rheumatoid arthritis; uveitis; AIDS; allergy;
 KW acquired immunodeficiency disease; neurological disorder; epilepsy;
 KW Alzheimer's disease; dementia; mental retardation; gastrointestinal;
 KW Parkinson's disease; ulcer; cirrhosis; reproductive; infertility;
 KW endometriosis; endocrine disorder; hyperparathyroidism; hyperlipidemia;
 KW hypercholesterolemia; hypoglycaemia; obesity; Reiter's syndrome;
 KW connective tissue disorder; osteoporosis; infection.
 XX Homo sapiens.
 XX WO200279473-A2.
 XX PD 10-OCT-2002.
 XX 09-JAN-2002; 2002WO-US01009.
 XX 12-JAN-2001; 2001US-261622P.
 PR 16-JAN-2001; 2001US-261864P.
 PR 16-JAN-2001; 2001US-261865P.
 PR 17-JAN-2001; 2001US-262164P.
 PR 17-JAN-2001; 2001US-262207P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262209P.
 PR 17-JAN-2001; 2001US-262215P.
 PR 18-JAN-2001; 2001US-263102P.
 PR 19-JAN-2001; 2001US-263599P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263064P.

XX WPI; 2001-343950/36.
 DR P-PSDB; AAU03525.
 XX Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 XX
 PS Example 1; Figure 1; 43pp; English.
 XX AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV), and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 XX Sequence 1083 BP; 366 A; 199 C; 225 G; 293 T; 3 other;
 SQ

Query Match 91.3%; Score 963; DB 22; Length 1083;
 Best Local Similarity 94.5%; Pred. No. 5.9e-223;
 Matches 927; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 2 ATGGAAAGTAGTGAATAATTAGCTAGCACTGGAGAGGGTCTTATGGGTGTATTCNAA 60
 DB 1 ATGGAAAGTAGTGAATAATTAGCTAGCACTGGAGAGGGTCTTATGGGTGTATTCNAA 60
 QY 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTCTTAAATAAATTGGGAACTGAAGAT 120
 DB 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTCTTAAATAAATTGGGAACTGAAGAT 120
 QY 121 GATCTCTGTGTAGAAATAGCACTAGAGAAATAGTAGTGTGAAGCAATTAACACAT 180
 DB 121 GATCTCTGTGTAGAAATAGCACTAGAGAAATAGTAGTGTGAAGCAATTAACACAT 180
 QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 240
 DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 240
 QY 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGTGTGAT 300
 DB 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGTGTGAT 300
 QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTTTTAACTTCTGTCTATACAT 360
 DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTTTTAACTTCTGTCTATACAT 360
 QY 361 AACTGTATTTCACAGAGATATTAACCTGAAATATTTCTTAATACTAAGCAAGGAATATC 420
 DB 361 AACTGTATTTCACAGAGATATTAACCTGAAATATTTCTTAATACTAAGCAAGGAATATC 420
 QY 421 AAGATTGTGACTTCGGGTTCGCAAAATTCGTA----- 454
 DB 421 AAGATTGTGACTTCGGGTTCGCAAAATTCGTAATTCACATGTTGGAGAAATTCAC 480
 QY 455 -----TTCCAGAGATGCTACACCGATTATGTAGCT 486
 DB 481 TTACCTGATCTGATTGATGATGCTTTCAGTTCACAGAGATGCTTACACCGATTATGTAGCT 540
 QY 487 ACGAGATGATACCGAGCTCTCGAACTTTCTGGAGATACCTAGTATGTTCTTCAGTC 546
 DB 541 ACGAGATGATACCGAGCTCTCGAACTTTCTGGAGATACCTAGTATGTTCTTCAGTC 600

QY 547 GATATATGGGCTATTGGTGTGTTTTCGAGAGCTCTGACAGGCGCAGCAGCTGTGGCCT 606
 DB 601 GATATATGGGCTATTGGTGTGTTTTCGAGAGCTCTGACAGGCGCAGCAGCTGTGGCCT 660
 QY 607 GGAAATACAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGGAAATTTAAATCCCA 666
 DB 661 GGAAATACAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGGAAATTTAAATCCCA 720
 QY 667 AGACATCAATCAATCTTTAAAGTAAACGGGTTTTCATGGGATCAGTATACCTGAGCCA 726
 DB 721 AGACATCAATCAATCTTTAAAGTAAACGGGTTTTCATGGGATCAGTATACCTGAGCCA 780
 QY 727 GAAGACATGGAACACTCTGAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACTTC 786
 DB 781 GAAGACATGGAACACTCTGAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACTTC 840
 QY 787 ATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCTCAACTCTCTGGAG 846
 DB 841 ATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCTCAACTCTCTGGAG 900
 QY 847 AGCTCCTACTTCTGATCTTTTCAAGAGGCGCCCAATTTAAGAAAGACAGCTATGAAGGA 906
 DB 901 AGCTCCTACTTCTGATCTTTTCAAGAGGCGCCCAATTTAAGAAAGACAGCTATGAAGGA 960
 QY 907 AGAAACAGAAAGCGCCCAACAG 927
 DB 961 AGAAACAGAAAGCGCCCAACAG 981

RESJUT 6
 AAD03816
 ID AAD03816 standard; cDNA; 1041 BP.
 XX
 AC AAD03816;
 XX
 DT 19-JUN-2001 (first entry)
 XX
 DE Human kinase cDNA #5.
 XX
 KW Human; kinase; gene therapy; bioreactor; mental disorder;
 KW biological disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 1..1041
 FT /*tag= a
 FT /product= "Human kinase #5"
 FT /note= "The coding region does not include stop codon"
 FT /partial
 XX
 PN WO200123579-A1.
 XX
 PD 05-APR-2001.
 XX
 PF 27-SEP-2000; 2000WO-US26621.
 XX
 PR 28-SEP-1999; 99US-0156511.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX
 DR WPI; 2001-266166/27.
 DR P-PSDB; AAE00494.
 XX
 PT New isolated human kinase polynucleotide useful for generating for
 PT antibodies, as reagents in diagnostic assays and for screening for
 PT compounds useful for treating mental, biological or medical diseases
 XX
 PS Claim 1; Page 31-32; 38pp; English.
 XX

RESULT 7	
AA03813	
ID	AA03813 standard; cDNA; 1068 BP.
XX	AC
XX	AA03813;
XX	
DT	19-JUN-2001 (first entry)
XX	
DE	Human kinase cDNA #2.
XX	
KW	Human; kinase; gene therapy; bioreactor; mental disorder;
KW	biological disorder; ss.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
PH	1..1068
FT	/*tag= a
FT	/product= "Human kinase #2"
FT	/note= "The coding region does not include stop codon"
FT	/partial
XX	
PN	WO200123579-A1.
XX	
PD	05-APR-2001.
XX	
PF	27-SEP-2000; 2000WO-US26621.
XX	
PR	28-SEP-1999; 99US-0156511.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Doroho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT
XX	
DR	WPI; 2001-266166/27.
DR	P-PSDS; AAE00491.
XX	
PT	New isolated human kinase polynucleotide useful for generating
PT	antibodies, as reagents in diagnostic assays and for screening for
PT	compounds useful for treating mental, biological or medical diseases
XX	
PS	Disclosure; Page 28; 38pp; English.
XX	

CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;

Query Match 87.1%; Score 823; DB 22; Length 1068;

Best Local Similarity 90.6%; Pred No. 3.8e-212;

Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 ATGGAAGTATGAAAAATAGCTAAGACTGGAGAGGCTTTATGGGTTGTATTCAAA 60
DB 1 ATGGAAGTATGAAAAATAGCTAAGACTGGAGAGGCTTTATGGGTTGTATTCAAA 60
QY 61 TGCAGAAACAAACCTCGGACAGTACTGTATTAAGAAATTTGGGATCTGAGAT 120
DB 61 TGCAGAAACAAACCTCGGACAGTACTGTATTAAGAAATTTGGGATCTGAGAT 120
QY 121 GATCTGTTGTTAAGAAATAGCATAAGAGAAATACGTATGTTGAAGCAATTAAGACAT 180
DB 121 GATCTGTTGTTAAGAAATAGCATAAGAGAAATACGTATGTTGAAGCAATTAAGACAT 180
QY 181 CCAATCTGTGAACTCATCGAGGTGTTGAGGAGAAAGAAATGCAATTTAGTATTTT 240
DB 181 CCAATCTGTGAACTCATCGAGGTGTTGAGGAGAAAGAAATGCAATTTAGTATTTT 240
QY 241 GAATCTGTGATCATACTTTTAATGAGCTGGAAGAAACCCCAATGGAAGTTCGTAT 300
DB 241 GAATCTGTGATCATACTTTTAATGAGCTGGAAGAAACCCCAATGGAAGTTCGTAT 300
QY 301 GGAGTGATCAAAAGCGTATTAAGCAAAACACITTCAGCTTTAATTTCTGTCATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTAAGCAAAACACITTCAGCTTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTACAGAGATATAAAGCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
DB 361 AACTGTATTACAGAGATATAAAGCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
QY 421 AAGATTTGTGACTTCGGGTTTGCAAAATTTCTGATTCAGGAGATGCTTACACCGATTAT 480
DB 421 AAGATTTGTGACTTCGGGTTTGCAAAATTTCTGATTCAGGAGATGCTTACACCGATTAT 480
QY 481 GTAGCTACGAGATGTTACGAGCTCTGAACTCTGTTGGGAGATCTCAGTATGTTCT 540
DB 481 GTAGCTACGAGATGTTACGAGCTCTGAACTCTGTTGGGAGATCTCAGTATGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTTGGTTGTTTTTGGCAGAGCTCCTGACAGGCGCAGCCTG 600
DB 541 TCAGTCGATATATGGGCTATTTGGTTGTTTTTGGCAGAGCTCCTGACAGGCGCAGCCTG 600
QY 601 TGGCCTGGAAATACAGATGGGACCACTTTTCTGATTAATCAGAACT----- 650
DB 601 TGGCCTGGAAATACAGATGGGACCACTTTTCTGATTAATCAGAACTTAGTAGAGACG 660

QY 651 ----- 650
DB 661 GGGTTTCCCATGTTGACAGGCTGTTCTGAACTCTTGACGTCAAGTGATCCACCTGCC 720
QY 661 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
DB 721 GTAGCTCTCAAGTCTGGAATTAACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
QY 695 AAAAGTAAACGGTTTTTCCATGGCATCAGATACCTTGAGCCAGACATGGAACTCTT 744
DB 781 AAAAGTAAACGGTTTTTCCATGGCATCAGATACCTTGAGCCAGACATGGAACTCTT 840
QY 745 GAGGAAAGTTCTCAGATGTTTCATCTGTCGCTCTGAACCTTCATGAAGGGTCTCTGAAG 804
DB 841 GAGGAAAGTTCTCAGATGTTTCATCTGTCGCTCTGAACCTTCATGAAGGGTCTCTGAAG 900
QY 805 ATGAATCCAGATGACAGATTAACCTGTTTCCAACTCCTGGAGAGCTCTACTTTGATCT 864
DB 901 ATGAATCCAGATGACAGATTAACCTGTTTCCAACTCCTGGAGAGCTCTACTTTGATCT 960
QY 865 TTCAAGAGGCGCCAAATTAAGAAAGACGCTAATGAGGAAAGACAGACGCCAA 924
DB 961 TTCAAGAGGCGCCAAATTAAGAAAGACGCTAATGAGGAAAGACAGACGCCAA 1020
QY 925 CAGGT 929
DB 1021 CAGGT 1025
RESULT 8
AA164248
ID AA164248 standard; cDNA: 882 BP.
XX
AC AA164248;
XX
DT 08-MAR-2002 (first entry)
XX
DE Human Kinase 14257 cDNA.
XX
KW Protein kinase; enzyme; cytostatic; osteopathic; hepatotropic;
KW antidiabetic; neuroprotective; antiarthritic; dermatological;
KW immunosuppressive; antiinflammatory; antithyroid; antiporiatic;
KW ophthalmological; antiallergic; antisthmatic; antidiatherosclerotic;
KW hypotensive; vasotropic; antiarrhythmic; virucide; anorectic;
KW metabolic; immunomodulator; analgesic; cellular proliferative disorder;
KW cancer; acute lymphoblastic leukaemia; Hodgkin's disease;
KW bone metabolism disorder; osteoporosis; immune system disorder;
KW inflammatory; diabetes mellitus; osteoarthritis; asthma;
KW cardiovascular disorder; hypertension; coronary artery disease;
KW endothelial cell disorder; psoriasis; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT 5'UTR 1..128
FT /*tag= a
FT CDS 129..815
FT /*tag= b
FT FT /product= "Protein kinase 14275"
FT 3'UTR 816..882
FT /*tag= c
XX
PN WO200179488-A2.
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12188.
XX
PR 13-APR-2000; 2000US-196910P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX

PI Kapeller-libermann R;
 XX WPI: 2002-034355/04.
 DR P-PSDB; AAG78547.
 XX
 PT New 14257 polypeptides (protein kinases), useful as diagnostic targets
 PT and therapeutic agents for controlling cellular proliferative and/or
 PT differentiative disorder, bone disorders, immune disorders and
 PT cardiovascular disorders -
 XX
 XX
 PS Claim 2: Fig 1A; 98pp; English.
 XX
 CC The invention relates to an isolated 14257 polypeptide and nucleic
 CC acid encoding it. The 14257 protein is a protein kinase that acts as a
 CC modulating agent in regulating a variety of cellular processes.
 CC including cell proliferation, differentiation, growth and division.
 CC The activity of the protein of the invention may be described as:
 CC cytosolic; osteopathic; hepatotropic; antidiabetic; neuroprotective;
 CC antiarthritic; dermatological; immunosuppressive; antiinflammatory;
 CC antithyroid; antipsoriatic; ophthalmological; antiallergic;
 CC antiasthmatic; antiatherosclerotic; hypotensive; vasotropic;
 CC antiarrhythmic; virucide; anorectic; metabolic; immunomodulator and
 CC analgesic. The protein of the invention may act as a novel diagnostic
 CC target or therapeutic agent controlling certain disorders, for example
 CC kinase-associated or other 14257-associated disorders. These may include
 CC cellular proliferative disorders such as cancers e.g. acute lymphoblastic
 CC leukemia or Hodgkin's disease. Other disorders include bone metabolism
 CC disorders such as osteoporosis. Disorders of the immune system, e.g.
 CC inflammatory, diabetes mellitus, osteoarthritis and asthma. Proteins of
 CC the invention may also be of use as therapeutic agents in cardiovascular
 CC disorders such as hypertension and coronary artery disease, and some
 CC endothelial cell disorders, including psoriasis. The current
 CC sequence represents a human kinase 14257 cDNA.
 XX
 SQ Sequence 882 BP; 288 A; 150 C; 198 G; 245 T; 1 other;
 XX
 Query Match 68.3%; Score 645.4; DB 24; Length 882;
 Best Local Similarity 99.1%; Pred. No. 3.3e-164;
 Matches 649; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 ATGGAAAGTATGAAATATAGCTAAGACTGAGAGGGTCTTATGGGGTTGATTTCAA 60
 DB 129 ATGGGAAGTATGAAATATAGCTAAGACTGAGAGGGTCTTATGGGGTTGATTTCAA 188
 QY 61 TCGAGAACAAACCTCGACAGTACTGCTGTAAATAATTTGGCAATCTGAAGAT 120
 DB 189 TCGAGAACAAACCTCGACAGTACTGCTGTAAATAATTTGGCAATCTGAAGAT 248
 QY 121 GATCCTGTTTGAAGAAATAGCAGTAAAGAAATACGTATGTTCAAGCAATTAAACAT 180
 DB 249 GATCCTGTTTGAAGAAATAGCAGTAAAGAAATACGTATGTTCAAGCAATTAAACAT 308
 QY 181 CCAATCTTTGGAACCTCATCGAGTGTTCAGAGAAAGAAAGAAATGCAATTTAGTTT 240
 DB 309 CCAATCTTTGGAACCTCATCGAGTGTTCAGAGAAAGAAAGAAATGCAATTTAGTTT 368
 QY 241 GAATACGTGTATACATCTTTTAAATGAGTGGAAAGAAACCCAAATGGAGTGGTAT 300
 DB 369 GAATACGTGTATACATCTTTTAAATGAGTGGAAAGAAACCCAAATGGAGTGGTAT 428
 QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
 DB 429 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 488
 QY 361 AACTGTATTTCAGAGATATTAACCTGAAATATTTCTTAATCTAAGCAGGAAATATC 420
 DB 489 AACTGTATTTCAGAGATATTAACCTGAAATATTTCTTAATCTAAGCAGGAAATATC 548
 QY 421 AAGATTGTGACTTCGGGTTTCACAAATTTCTGATTCAGGAGATGGCTACACCGATTAT 480
 DB 549 AAGATTGTGACTTCGGGTTTCACAAATTTCTGATTCAGGAGATGGCTACACCGATTAT 608
 QY 481 GTAGCTACAGATGGTACCGAGCTCTGAACTCTTTGTGGGAGATACCTAGCTATGGTCT 542

Db 609 CTAGCTACGAGTGGTACCGAGCTCTGAACTTCTTGTGGAGATACCTAGTGTCT 668
 QY 541 TCAGTCGATATATGGCTATTTGGTGTGTTTTCAGAGCTCTTGACAGCCAGCCACTG 600
 Db 669 TCAGTCGATATATGGCTATTTGGTGTGTTTTCAGAGCTCTTGACAGCCAGCCACTG 728
 QY 601 TGGCTCTGAAATACAGATGTGGACCACTTTATCTGATATATCAGAACACTAGGAA 655
 Db 729 TGGCTCTGAAATACAGATGTGGACCACTTTATCTGATATATCAGAACACTAGGTA 783

RESULT 9
 AAF44669
 ID AAF44669 standard; cDNA; 911 BP.
 XX
 AC AAF44669;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase cDNA, SEQ ID NO: 49.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200073469-A2.
 XX
 PD 07-DEC-2000.
 XX
 XX 26-MAY-2000; 2000WO-US14842.
 XX
 PR 28-MAY-1999; 99US-0136503.
 XX
 XX (SUGB-) SUGEN INC.
 XX
 PA Plowman GD, Martinez R, Whyte D, Sudersanam S;
 P: WPI: 2001-032161/04.
 XX
 DR P-PSDB; AAB56642.
 XX
 PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 PS
 XX Example 4; Fig 2; 310pp; English.
 CC The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 SQ Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;
 Query Match 68.0%; Score 643; DB 22; Length 911;

Best Local Similarity 100.3%; Pred. No. 1,58-163;	
Matches	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	285 AAATGGAGTGTGCTGATGGAGTGATCAAAAGCGTATATGCGAAACACTTTCAAGCTCTTAA 344
DB	1 AAATGGAGTGTGCTGATGGAGTGATCAAAAGCGTATATGCGAAACACTTTCAAGCTCTTAA 60
QY	345 TTTCTGTATATACATAACTGTATTTCACAGNATATATAAACCTGTAATAATTTCTTAATAAC 404
DB	61 TTTCTGTATATACATAACTGTATTTCACAGNATATATAAACCTGTAATAATTTCTTAATAAC 120
QY	405 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCACAGGAGA 464
DB	121 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCACAGGAGA 180
QY	465 TGCCTACACCGATATGTAGCTACGAGATGTCACCGAGCTCTGAACTTCTTTGTGGGAGA 524
DB	181 TGCCTACACCGATATGTAGCTACGAGATGTCACCGAGCTCTGAACTTCTTTGTGGGAGA 240
QY	525 TACTCAGTATGGTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCACAGAGCTCT 584
DB	241 TACTCAGTATGGTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCACAGAGCTCT 300
QY	585 GACGSCCAGCCACTGTGGCTCGAAATCAGATGTGGACCACTTTTCTCATTAATCAG 644
DB	301 GACGSCCAGCCACTGTGGCTCGAAATCAGATGTGGACCACTTTTCTCATTAATCAG 360
QY	645 AACCTAGGAAAAATTAATCCCAAGACATCAATCAATCTTTTAAAGTAACCGGTTTTTCCA 704
DB	361 AACCTAGGAAAAATTAATCCCAAGACATCAATCAATCTTTTAAAGTAACCGGTTTTTCCA 420
QY	705 TGGCATCAGTATACCTGAGCCAGAGACATGGAACTCTTGGGGAAGAGTCTTCAGATGT 764
DB	421 TGGCATCAGTATACCTGAGCCAGAGACATGGAACTCTTGGGGAAGAGTCTTCAGATGT 480
QY	765 TCATCCTGTGGCTGTGAATCTTATGAAGGGGTGCTSAAGATGAATCCAGATTCACAGATT 824
DB	481 TCATCCTGTGGCTGTGAATCTTATGAAGGGGTGCTSAAGATGAATCCAGATTCACAGATT 540
QY	825 AACCTGTCTCCAACTCTCGAGAGCTCTACTTTGATCTTTTTCAGAGGCCCCAATTTAA 884
DB	541 AACCTGTCTCCAACTCTCGAGAGCTCTACTTTGATCTTTTTCAGAGGCCCCAATTTAA 600
QY	885 AAGAAAAGCAGCTAATGAAGGAGAAAAAGAGACGCCAACAG 927
DB	601 AAGAAAAGCAGCTAATGAAGGAGAAAAAGAGACGCCAACAG 643

```

PR 28-MAY-1999; 99US-0136503.
XX (SUGEN-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
XX WF; 2001-032161/04.
XX P-PSDB; AAB65643.
XX
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
XX PT treating immune-related diseases and disorders, cardiovascular disease,
XX PT neurodegenerative diseases and/or cancers -
XX
XX Disclosure; Fig 2; 30pp; English.
XX
XX The present sequence encodes a novel protein kinase. The nucleic acids
XX CC and the protein kinases they encode may be used in the treatment and
XX CC diagnosis of diseases associated with inappropriate kinase expression
XX CC such as immune-related diseases and disorders, cardiovascular disease,
XX CC neurodegenerative diseases and/or cancers. The nucleic acids and
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays. The kinase polypeptides may be used as antigens in the production
XX CC of antibodies of kinase expression and activity. Anti-kinase antibodies
XX CC and kinase antagonists may also be used to down regulate kinase
XX CC expression and activity. Diseases related to kinase expression and
XX CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
XX CC disorders, complications of organ transplantation, myocardial infarction,
XX CC immune disorders, cardiomyopathies, strokes, renal failure,
XX CC oxidative-stress related disorders, chronic inflammatory bowel disease,
XX CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
XX CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
XX CC reproductive disorders.
XX
XX Sequence 26-5 Bp; 742 A; 580 C; 585 G; 708 T; 0 other;
XX
Query Match 59.2%; Score 559.2; DB 22; Length 2615;
Best Local Similarity 81.8%; Pred. No. 9,7e-141;
Matches 658; Conservative 0; Mismatches 143; Indels 3; Gaps 1;
XX
QY 135 GAAATATGACACTAAGAGAATACTGTGTGAAGCAATTAACCATCCAAATCTGTGAA 194
Db 1 GAAATATGCCCTCGGGAATCCGTATGCTGAAG---TTGAACACCCAAACCTGTGAA 57
QY 195 CTTATCGAGGTGTTTCAGGAGAAAAAGGAAATGCAATTTAGTTTTTGAATCTGTGATCA 254
Db 58 CCTCATCGAGGTGTTTCAGAGAAAGAGAAAGATGCACTAGTTTTTGAGTACTGTGATCA 117
QY 255 TACACTTTTAAATGAGCTCGGAAGAAACCCAAATGCAGTTGCTGTGGAGTGATCAAAAG 314
Db 1-8 CACACTGTTAAACGAGCTGGAGAAACCCAAACGGAGTTTCTGATGGAGTGATTAAG 177
QY 315 CGTATTATGGCAAACTTCAAGCTCTTAATTTCTGTATATACATACTGATTCACAG 374
Db 178 TGTGCTATGGCAACCTTCAAGCCCTTAACCTTCTGTCAACAGCAATTTGATTCATCG 237
QY 375 AGATATAAACCTGGAATATTTCTATTAAGCAAGGAATTAATCAAGATTTGTGACTT 434
Db 238 GGAATATAAACCTTGAAGACATCCTAATAACCAAGCAAGGATGATTAAGATTTGTGACTT 297
QY 435 CGGGTTTGCACAAATTTCTGATTCAGGAGATGCTTACACCGGATTATGTAGCTACGAGATG 494
Db 298 TGGATTTGCACGAATTTCTAATTCAGGAGACGCTTACACAGACTATGTTGCCACAGGTG 357
QY 495 GTACCGAGCTCTGAACTTCTTGTGGGAGATACTCAGTATGTTTCTTCAGTGCATATATG 554
Db 358 GTACCGAGCCCCGAACCTTCTCGTGGGAGACAGAAAGTACGGTTCCTCTGTAGACGTGTG 417
QY 555 GGCATTGTGTGTTTTTGCAGAGCTCTCTGACAGCCAGCCAGCACTGTGCGCTGGAAATC 614
Db 418 GGCCGTGCGCTGTGTTTTTGCAGAGCTCTCTGACGGGTACGCCACTCTGCGCGGGAATC 477
QY 615 AGATGTGGACCAACTTTATCTGATTAATCAGAACTAGGAAATTTAATCCCAAGCATCA 674

```

RESULT 10	
AAF44670	AAF44670
XX	AAF44670 standard; cDNA; 2615 bp.
XX	
XX	AAF44670;
XX	
XX	27-MAR-2001 (first entry)
XX	
XX	Novel protein kinase cDNA, SEQ ID NO: 50.
XX	
KW	Human; mouse; protein kinase; antiarthritic; anisclerotic; osteopathic;
KW	immunosuppressive; cardiac; renal; antiinflammatory; antiaschmatic;
KW	dermatological; antidiabetic; antifertility; gene therapy; vaccine;
KW	immune disorder; cardiovascular disease; neurodegenerative disease;
KW	cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW	inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
OS	Musculus.
XX	
XX	WO200073469-A2.
PN	
XX	
PD	07-DEC-2000.
XX	
XX	26-MAY-2000; 2000WO-US14842.
PP	
XX	

Dd 478 CGACGTGACACAGCTTTACCTCATCATCAGGACGTTGGGAAGCTGATTCAGACACCA 537
Qy 675 ATCAATCTTTAAAGTAAGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGCAT 734
Dd 538 GTCATCTTTAGGATACCAAGTTTTTCGGGGGATCAGCATCTCAACGAGGACAT 597
Qy 735 GGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCTGTGGCTGTGAACCTCATGAAGGG 754
Dd 598 GCAGACTCTTGAAGAAAATCTCAATGTTTCAGCTGTGCTTTAAGTTTCATGAAGGG 657
Qy 795 GTGTGTGAGATGAATCCAGATGACAGATTAACTGTTTCCCACTCTGTGAGAGCTCTTA 854
Dd 658 ATGCGCTGAAGATGAATCTGTGATGAGAGCTGACCTGTGCCAGCTGTGACAGTGCCTA 717
Qy 855 CTTTGATCTTTTCAAGAGGCCCCAATTAAGAGAAAAGCAGCTTAATGAAGGAAGAAACAG 914
Dd 718 CTTTGAGTCTTTCAAGAGGATCAATGAAGAAAAGCCCGAGTGGGGAGAGCCG 777
Qy 915 AAGAGCCCAAGGTAATCTTCCGCT 938
Dd 778 AAGGGCCAGCAGAAATCAACTGCT 801

RESULT 11

ABX34679
ID ABX34679 standard; cDNA; 1281 BP.

AC ABX34679;
XX

DT 13-FEB-2003 (first entry);
XX

DE Human mdat cDNA SEQ ID 240.
XX

KW MDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
KW haemostatic; nephrotropic; antianemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis; gene; ss.

XX Homo sapiens.
OS

PN WO200279449-A2.
XX

PD 10-OCT-2002.
XX

PF 27-MAR-2002; 2002WO-US09944.
XX

PR 28-MAR-2001; 2001US-279619P.
XX

PR 29-MAR-2001; 2001US-280067P.
XX

PR 29-MAR-2001; 2001US-280068P.
XX

PR 16-MAY-2001; 2001US-291280P.
XX

PR 17-MAY-2001; 2001US-291829P.
XX

PR 19-JUN-2001; 2001US-299428P.
XX

PR 20-JUN-2001; 2001US-299776P.
XX

PR 20-JUN-2001; 2001US-300031P.
XX

PA (INCYTE GENOMICS INC.
XX

PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX

PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
XX

PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerslein EH;
XX

PI Zeralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX

PI Flores V, Marwaha R, Lo A, Jan RY, Urashka ME;
XX

XX WPI; 2003-058431/05.
DR

DR P-PSDB; ABU11689.
XX

XX New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers

PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
or hepatitis -
XX
XX Claim 1; SEQ ID NO 240; 339pp + Sequence Listing; English.
CC This invention describes a novel disease detection and treatment molecule
polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
antianemic, antipsoriatic and hepatotropic activity. The polynucleotides
and the polypeptides of the invention can be used for gene therapy, of
protein replacement therapy and are useful for treating a variety of
diseases or conditions. These polypeptides or polynucleotides are
particularly useful for diagnosing, treating or preventing cell
proliferative disorders (e.g. cancers including adenocarcinoma,
leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
syndrome, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
hepatitis. ABX3440-ABX34835 encode the MDDT polypeptides represented in
ABU11450-ABU11845, described in the disclosure of the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1281 BP; 428 A; 131 C; 317 G; 345 T; 0 other;

Query Match 57.5%; Score 543.2; DB 25; Length 1281;

Best Local Similarity 99.5%; Pred. No. 1.5e-136;

Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGAAAAGTATGAAAATTTAGCTAAGCACTGGAGAGGGCTTATGGGGTTGTATTCAAA 60

Dd 754 ATGGAAAAGTATGAAAATTTAGCTAAGCACTGGAGAGGGCTTATGGGGTTGTATTCAAA 763

Qy 61 TCGAAGAACAAAACCTCTGGCAAGTAGTAGCTGTTAAAAAATTTGTGAATCTGAAGAT 120

Dd 764 TCGAAGAACAAAACCTCTGGCAAGTAGTAGCTGTTAAAAAATTTGTGAATCTGAAGAT 823

Qy 121 GATCCTGTGTTAGAAAATAGCACTAAGAGAAATAGTATGTTGAAGCAATTAACAT 180

Dd 824 GATCCTGTGTTAGAAAATAGCACTAAGAGAAATAGTATGTTGAAGCAATTAACAT 883

Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAAGGAAAATGCAATTTAGTTTTT 240

Dd 884 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAAGGAAAATGCAATTTAGTTTTT 943

Qy 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300

Dd 944 GAATCTGTGATCATACACTTTTAAATGGGCTGGAAGAAACCCCAATGGAGTTGCTGAT 1003

Qy 301 GGAGTGATCAAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360

Dd 1004 GGAGTGATCAAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 1063

Qy 361 AACTGTATTCACAGAGATATAAAACCTGAAAAATTTCTAATACTAAGCAAGGAATAATC 420

Dd 1064 AACTGTATTCACAGAGATATAAAACCTGAAAAATTTCTAATACTAAGCAAGGAATAATC 1123

Qy 421 AAGATTTGTGACTTCGGGTTTGACAAATTTCTGATTCAGAGAGATGCTACACCGATTAT 480

Dd 1124 AAGATTTGTGACTTCGGGTTTGACAAATTTCTGATTCAGAGAGATGCTACACCGATTAT 1183

Qy 481 GTAGCTACGAGATGTTACCGAGCTCCTGAATCTTCTTGTGGGAGATACTCAGTATGGTTCT 540

Dd 1184 GTAGCTACGAGATGTTACCGAGCTCCTGAATCTTCTTGTGGGAGATACTCAGTATGGTTCT 1243

Qy 541 TCAGTCGA 548

Dd 1244 TCAGTCGA 1251

RESULT 12

ABX03818

ID AAD03818 standard; cDNA; 1819 BP.

XX AAD03818;
 XX AC
 XX D7
 XX DE 19-JUN-2001 (first entry)
 XX DE Human kinase cDNA #7.
 XX DE
 XX KW Human; kinase; gene therapy; bioreactor; mental disorder;
 XX KW biological disorder; polymorphism; ss.
 XX OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX PH 1..137
 XX FT 5'UTR /*tag= a
 XX FT CDS 138..701
 XX FT /*tag= b
 XX FT /*product= "Human kinase #1"
 XX FT 702..1819
 XX FT 3'UTR /*tag= c
 XX FT misc_difference 1684
 XX FT /*tag= d
 XX FT /*note= "polymorphism site; given as r in the sequence"
 XX
 XX WO200123579-Ai.
 XX PD
 XX PD 05-APR-2001.
 XX PD
 XX PD 27-SEP-2000; 2000WO-US26621.
 XX PD
 XX PD 28-SEP-1999; 99US-0156511.
 XX PD (LEXI-) LEXICON GENETICS INC.
 XX PD
 XX PD Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX PD WPI; 2001-266166/27.
 XX PD P-PSDB; AAE00490.
 XX PD
 XX PD New isolated human kinase polynucleotide useful for generating
 XX PD antibodies, as reagents in diagnostic assays and for screening for
 XX PD compounds useful for treating mental, biological or medical diseases -
 XX PD
 XX PS Disclosure; Page 34-35; 38pp; English.
 XX
 XX The present sequence encodes a novel human protein (NHP) which
 XX is a human kinase. A polymorphism was identified in the 3' UTR
 XX of the present sequence. The human kinases share structural
 XX similarity with animal kinases, more particularly serine or
 XX threonine protein kinases. Human kinase cDNA is useful for the
 XX detection of mutant human kinase for the diagnosis of disease,
 XX and also as a therapeutic. It is useful for screening drugs
 XX effective in the treatment of symptomatic or phenotypic
 XX manifestations perturbing the normal function of NHP in the
 XX body. The NHP nucleotide sequences are useful for generation of
 XX antibodies, as reagents in diagnostic assays, for the
 XX identification of other cellular gene products related to human
 XX kinases, and as reagents in assays for screening compounds that
 XX are useful for treating mental, biological or medical disorders.
 XX NHP oligonucleotides are used as probes. The labeled NHP probes
 XX are useful for screening human genomic library for identifying
 XX polymorphisms and as primers in amplification assays to detect
 XX mutations within the exons, introns and splice sites that can
 XX be used in diagnostics and pharmacogenomics. Nucleotide construct
 XX encoding NHP products are used to genetically engineer cells
 XX in vivo that functions as bioreactors in the body delivering a
 XX continuous supply of NHP to the body. Nucleotide constructs
 XX encoding functional NHPs are used in gene therapy for the
 XX modulation of NHP expression.
 XX
 XX Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;
 XX Query Match 53.7%; Score 507; DB 22; Length 1819;

Best Local Similarity 69.8%, Pred. No. 1.le-126;
 Matches 929; Conservative 0; Mismatches 0; Indels 402; Gaps 2;
 QY 1 ATGGAAGATGAAATAGCTAGCTAGAGAGGGTCTTATGGGTTGTATTCAAA 60
 DB |||||
 QY 138 ATGGAAGATGAAATAGCTAGCTAGAGAGGGTCTTATGGGTTGTATTCAAA 197
 DB |||||
 QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTTTAAAGAAATTTTGGAAATCTGAAGAT 120
 DB |||||
 QY 121 GATCCTGTGTTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
 DB |||||
 QY 258 GATCCTGTGTTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 317
 DB |||||
 QY 181 CCAATCTTGTGAACCTCATCGAGGTCTTCAGGAGAAAGGAAATGCATTTAGTTTTT 240
 DB |||||
 QY 318 CCAATCTTGTGAACCTCATCGAGGTCTTCAGGAGAAAGGAAATGCATTTAGTTTTT 377
 DB |||||
 QY 241 GAATACCTGATCATACACTTTTAAATGAGCTGAAAGAAACCCAAATGGAGTTGCTGAT 300
 DB |||||
 QY 378 GAATACCTGATCATACACTTTTAAATGAGCTGAAAGAAACCCAAATGGAGTTGCTGAT 437
 DB |||||
 QY 301 GGAGTGATCAAAACGCTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
 DB |||||
 QY 438 GGAGTGATCAAAACGCTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT 497
 DB |||||
 QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAAGCAAGGAATATC 420
 DB |||||
 QY 438 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAAGCAAGGAATATC 557
 DB |||||
 QY 421 AAGATTTGTACTTCGGGTTTGCAAAATTCG ----- 453
 DB |||||
 QY 558 AAGATTTGTACTTCGGGTTTGCAAAATTCGAGTTGGACTTCATCTTCTCTGTGCC 617
 DB |||||
 QY 454 ----- 453
 DB 618 TCCTTGATTGCTTAATAGTTGACCTTCTGAATCTTTTCTGCCAATTCAGAGATTTT 677
 DB |||||
 QY 454 ----- 453
 DB 678 CTCCTGGCTTGGATCCATTGCTGACACAGTGTTCACCATGGGCCCCAGGCTCATCTCGA 737
 DB |||||
 QY 454 ----- 453
 DB 738 ACTCTGSCCTCAAGTGAATCTTCCACCTCGGCTCCCAAGTGTGATGCAAGTGTG 797
 DB |||||
 QY 454 ----- 453
 DB 738 AGCCACCGTGGCCAGCAGATTTTTCAAACAATAACTACTGAGAGCTCACAAGATTTGTT 857
 DB |||||
 QY 454 -----ATTCCAGAGATGCTACACC 474
 DB 858 TTATGGGAAACAAATTCGAAACAAATTTCTTGAAGACGCATTCAGGAGATGCTACACC 917
 DB |||||
 QY 475 GATTATGTAGCTAGAGATGTCACCGAGCTCCTGAACCTCTTGTGGAGATATCTCAGTAT 534
 DB |||||
 QY 918 GATTATGTAGCTAGAGATGTCACCGAGCTCCTGAACCTCTTGTGGAGATATCTCAGTAT 977
 DB |||||
 QY 535 GGTTCCTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 594
 DB |||||
 QY 978 GGTTCCTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 1037
 DB |||||
 QY 595 CCATGTGGCTCGAAATCAGATGTGAGCAACTTTATCTGATAATCAGAACT----- 650
 DB 1038 CCATGTGGCTCGAAATCAGATGTGAGCAACTTTATCTGATAATCAGAACTAGTA 1097
 DB |||||
 QY 651 ----- 650
 DB 1098 GAGAGGGGTTTCCCATGTTGACAGGCTGGTCTCGAACTTTGACGTCAAGTGATCCA 1157
 DB |||||
 QY 651 -----AGGAAATTAATCCCAAGACATCAATCA 678
 DB |||||

Db 1158 CTGCGGTAGGCTCTCAAGAGTCTGGAATTACAGAAATTAATCCCAAGACATCAATCA 1217
 Qy 679 ATCTTTAAAGTAACGGGTTTTCATGGCATCATGATATACCTGAGCCAGAGACATGGA 738
 Db 1218 ATCTTTAAAGTAACGGGTTTTCATGGCATCATGATATACCTGAGCCAGAGACATGGA 1277
 Qy 739 ACTCTTGAGGAAAGTCTTCAGATGTTTCATCTCTGGCTCTGAACTTCATCAAGGGGTGT 798
 Db 1278 ACTCTTGAGGAAAGTCTTCAGATGTTTCATCTCTGGCTCTGAACTTCATCAAGGGGTGT 1337
 Qy 799 CTGAAGATGATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAGAGCTCTACTTT 858
 Db 1338 CTGAAGATGATCCAGATGACAGATTAACTGTTCCCACTCTCTGGAGAGCTCTACTTT 1397
 Qy 859 GATTCCTTTCAAGAGGCCCAATTAAGGAAAGACGTAATGAAGGAACACAGAGA 918
 Db 1398 GATTCCTTTCAAGAGGCCCAATTAAGGAAAGACGTAATGAAGGAACACAGAGA 1457
 Qy 919 GCGCAACAGGT 929
 Db 1458 GCGCAACAGGT 1468

RESULT 13

ABZ77165
 ID ABZ77165 standard; cDNA; 1086 BP.
 AC
 XX
 DT
 XX
 DE Human protein kinase encoding cDNA SEQ ID NO:79.
 XX
 KW Human; protein kinase; enzyme; antiasthmatic; antiinflammatory;
 KW antidiabetic; antiparkinsonian; antimigraine; cardiant; cytosstatic;
 KW immunosuppressive; vulnerary; gene therapy; COPD; asthma; migraine;
 KW chronic obstructive pulmonary disease; non-insulin dependent diabetes;
 KW Parkinson's disease; myocardial infarction; inflammatory bowel disease;
 KW autoimmune disorder; allograft rejection; graft versus host disease;
 KW cancer; leukaemia; wound granulation; gene; ss.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..1086
 /*tag= a
 /partial
 /product= "protein kinase"
 /note= "no start or stop codons given"

WO2003000901-A2.

03-JAN-2003.

24-JUN-2002; 2002WO-IB02358.

26-JUN-2001; 2001US-301098P.

26-NOV-2001; 2001US-332870P.

(DECO-) DECODE GENETICS EHF.

Martinez RAM, Sigurdsson GT;

WPI: 2003-201429/19.

P-PSDB; ABP96087.

New protein kinase genes and polypeptides, useful for diagnosing
 PT diseases associated with a protein kinase, or in gene therapy for
 PT treating e.g. Parkinson's disease, migraine, myocardial infarction,
 PT allograft rejection or cancers -

Claim 1; Page 89; 258pp; English.

CC ABZ77126 to ABZ77165 encode the human protein kinases given in ABP96048
 CC to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
 CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cytosstatic,
 CC immunosuppressive and vulnerary activities, and can be used in gene
 CC therapy. A protein kinase therapeutic agent from the present invention,
 CC particularly a protein kinase gene agonist or antagonist, can be used
 CC for treating a disease or condition associated with a protein kinase in
 CC an individual. These diseases include chronic obstructive pulmonary
 CC disease (COPD), asthma, non-insulin dependent diabetes, Parkinson's
 CC disease, migraine, myocardial infarction, inflammatory bowel disease,
 CC autoimmune disorders (e.g. allograft rejection or graft vs. host
 CC disease), cancers (e.g. leukaemia) or wound granulation.
 XX
 SQ Sequence 1086 BP; 356 A; 227 C; 210 G; 293 T; 0 other;
 Query Match 53.0%; Score 501; DB 25; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 3.6e-125; Indels 0; Gaps 0;
 Matches 501; Conservative 0; Mismatches 0;
 Qy 252 GTTCTGATGGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGT 351
 Db 283 GTTCTGATGGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGT 342
 Qy 352 CATATACATACTCTATTCACAGAGATATAAACCCTGAAATATTTCTAATACTAAGCAA 411
 Db 343 CATATACATACTCTATTCACAGAGATATAAACCCTGAAATATTTCTAATACTAAGCAA 402
 Qy 412 GGAATAATCAAGATTGTGACTTCGGGTTTGCAAAATTTCTGATTCAGGAGATGCTTAC 471
 Db 403 GGAATAATCAAGATTGTGACTTCGGGTTTGCAAAATTTCTGATTCAGGAGATGCTTAC 462
 Qy 472 ACCGATATGTAGCTACAGATGGTACCGAGCTCTCTGAACTTCTTGTGGGAGATCTCAG 531
 Db 463 ACCGATATGTAGCTACAGATGGTACCGAGCTCTCTGAACTTCTTGTGGGAGATCTCAG 522
 Qy 532 TATGTTCTTCAGTCGATATATGGGCTATTGGTGTGTTTTGCGAGAGCTCTGACAGGC 591
 Db 523 TATGTTCTTCAGTCGATATATGGGCTATTGGTGTGTTTTGCGAGAGCTCTGACAGGC 582
 Qy 592 CAGCACCTGTGGCTGGAAATCAGATGTGACCAACTTTTCTGATAATCAGAACACTA 651
 Db 583 CAGCACCTGTGGCTGGAAATCAGATGTGACCAACTTTTCTGATAATCAGAACACTA 642
 Qy 652 GGAATAATTAATCCAGACATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATC 711
 Db 643 GGAATAATTAATCCAGACATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATC 702
 Qy 712 AGTATACCTGAGCCAGAGACATGGAACACTTTGAGGAAAGTTTCTCAGATGTTCTATCT 771
 Db 703 AGTATACCTGAGCCAGAGACATGGAACACTTTGAGGAAAGTTTCTCAGATGTTCTATCT 762
 Qy 772 GTGGCTCTGAACCTTCATGAAG 792
 Db 763 GTGGCTCTGAACCTTCATGAAG 783
 RESULT 14
 AAD03812
 C D AAD03812 standard; cDNA; 561 BP.
 XX
 AC AAD03812;
 XX
 DT 19-JUN-2001 (first entry)
 XX
 DE Human kinase cDNA #1.
 XX
 KW Human; kinase; gene therapy; bioreactor; mental disorder;
 KW biological disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS Location/Qualifiers
 1..561

CC are useful for screening human genomic library for identifying
 CC polymorphisms and as primers in amplification assays to detect
 CC mutations within the exons, introns and splice sites that can
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct
 CC encoding NHP products are used to genetically engineer cells
 CC in vivo that functions as bioreactors in the body delivering a
 CC continuous supply of NHP to the body. Nucleotide constructs
 CC encoding functional NHPs are used in gene therapy for the
 CC modulation of NHP expression.

XX
 XX Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;

QY	1	ATGGAAAGTATGAAATTAAGCTAGAGCTGAGAGGGCTCTATGGGTGTATTCAAA	60
Db	1	ATGGAAAGTATGAAATTAAGCTAGAGCTGAGAGGGCTCTATGGGTGTATTCAAA	60
QY	61	TGCAGAAACAAACCTCTGGACAGTACTAGCTGTAAATAATTTGGAACTCGAAGAT	120
Db	61	TGCAGAAACAAACCTCTGGACAGTACTAGCTGTAAATAATTTGGAACTCGAAGAT	120
QY	121	GATCCTGTGTTAAGAAATAGCCTAAGAGAAATACGTAATTTGAAGCAATTAACAT	180
Db	121	GATCCTGTGTTAAGAAATAGCCTAAGAGAAATACGTAATTTGAAGCAATTAACAT	180
QY	181	CCAAATCTTTGCAACCTCATCGAGGTGTCAGGAGAAAAGGAAAATCCATTAGTTT	240
Db	181	CCAAATCTTTGCAACCTCATCGAGGTGTCAGGAGAAAAGGAAAATCCATTAGTTT	240
QY	241	GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTCTGAT	300
Db	241	GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTCTGAT	300
QY	301	GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACAT	360
Db	301	GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACAT	360
QY	361	AACGTATTTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC	420
Db	361	AACGTATTTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC	420
QY	421	AAGATTGTGACTTCGGGTTTCACAAATTCCTGATT	456
Db	421	AAGATTGTGACTTCGGGTTTCACAAATTCCTGATT	456

Search completed: November 15, 2003, 16:54:33
 Job time : 301.918 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM: nucleic - nucleic search, using sw model

Run on: November 15, 2003, 16:54:37 ; Search time 67.0921 Seconds
(without alignments)
6216.928 Million cell updates/sec

Title: US-09-671-050-11

Perfect score: 945

Sequence: 1 atggaaaagatgaaaaaatt.....aggtacttcgcgcacaaagt 945

Scoring table: IDENTITY NUC

Gapop 10_0, Gapext 1_0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1:39956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCJS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	287	30.4	1701	US-09-411-628-5	Sequence 5, Appli
2	287	30.4	3080	US-09-411-628-3	Sequence 3, Appli
3	171.8	18.2	903	US-08-874-347-9	Sequence 9, Appli
4	171.8	18.2	903	US-09-093-522-9	Sequence 9, Appli
5	163.4	17.3	1476	US-08-969-106-1	Sequence 1, Appli
6	163.4	17.3	1476	US-09-338-125-1	Sequence 1, Appli
7	163.4	17.3	1476	US-09-266-225D-13	Sequence 13, Appli
8	163.4	17.3	1635	US-09-417-197-112	Sequence 112, App
9	163.4	17.3	1635	US-09-417-197-114	Sequence 114, App
10	155.6	16.5	1050	US-09-220-112-3	Sequence 3, Appli
11	136.6	14.5	1825	US-09-620-312D-313	Sequence 313, App
12	134.6	14.2	1089	US-08-154-915-1	Sequence 1, Appli
13	134.6	14.2	1089	US-08-464-517-37	Sequence 37, Appli
14	134.6	14.2	1089	US-08-246-361A-37	Sequence 37, Appli
15	134.6	14.2	1089	US-08-463-772-37	Sequence 37, Appli
16	134.6	14.2	1089	US-08-463-772-37	Sequence 37, Appli
17	131	13.9	1002	US-08-463-090B-3	Sequence 3, Appli
18	117.2	12.4	1070	US-08-463-090B-5	Sequence 5, Appli
19	112	11.9	1308	US-09-801-861-4	Sequence 4, Appli
20	112	11.9	2203	US-08-801-861-1	Sequence 1, Appli
21	103	10.9	2747	US-08-874-347-1	Sequence 1, Appli
22	103	10.9	2747	US-09-093-522-1	Sequence 1, Appli
23	102	10.8	2467	US-09-206-344A-3	Sequence 3, Appli
24	98	10.4	1673	US-09-347-801-15	Sequence 15, Appli
25	97.8	10.3	1100	US-08-950-449A-18	Sequence 18, Appli
26	97.8	10.3	3774	US-08-950-449A-11	Sequence 11, Appli
27	97.8	10.3	3813	US-08-469-421-11	Sequence 11, Appli

Sequence 11, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 1, Appli
Sequence 13, Appli
Sequence 13, Appli
Sequence 13, Appli
Sequence 13, Appli
Sequence 13, Appli
Sequence 13, Appli
Sequence 13, Appli
Sequence 13, Appli
Sequence 1, Appli
Sequence 64, Appli
Sequence 46, Appli
Sequence 1, Appli
Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-411-628-5
; Sequence 5, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; EARLIER FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-09-411-628-5

Query Match 30.4%; Score 287; DB 4; Length 1701;
Best Local Similarity 59.1%; Pred. No. 1.9e-71;
Matches 510; Conservative 0; Mismatches 350; Indels 3; Gaps 1;
QY 1 ATGGAAGATGATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60
DB 1 ATGGAAGATGATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60
QY 61 TGCAGAACAAACCTCTGGACAAGTAGTAGCTTTAAATAATTTTGGAAATCTGAAGT 120
DB 61 TGTAGGAATAAGATAGTGGAGAAATTTGTGCCATCAAGAGTTCTTAGAAGTGTATG 120
QY 121 GATCTCTGTTTGAAGAAATAGCACTAAGCAATATCTATGTTGAGCAATTTAAACAT 180
DB 121 GACAAATGGTTTAAATAATTTGCTATCGAGAAATCAAGTACTAAGCAACTGAGCAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAAGAGAAAAAGGAAATTCATTTAGTTTTT 240
DB 181 GAAATTTGGTGAATCTGTTGGAGGTGTGTAATAAAAAAAGGATGCTACCTAGTCTTT 240
QY 241 GAACTACTGTGATACATACACTTTTAAATGAGCTGGAGAGAACCCAAATCGAGTGTGAT 300
DB 241 GAACTACTGTGATACATACACTTTTAAATGAGCTGGAGAGAACCCAAATCGAGTGTGAT 300
QY 301 CGAGTGTCAAAAGCGTATTATGCAACCAACTCTCAAGCTCTTAATTTCTGTATATACAT 360
DB 301 CGAGTGTCAAAAGCGTATTATGCAACCAACTCTCAAGCTCTTAATTTCTGTATATACAT 360
QY 361 AACTGTATTACAGAGATATAAACCCTGAAATAATTTCTTAATACTAAGCAAGGAATAATC 420

```
Db 361 AATATCATACATAGAGATATAAAGCCAGAGAAATATATGGTCTCCAGTCTGCGGTGTC 420
Qy 421 AAGATTGTGACTTCGGGTTTGGCAAAATCTG---ATTCCAGAGATSCCTACACCGAT 477
Db 422 AAGTTATGTGATTTTGGATTGTGACGGACACTGGCACTCCCGAGAGGTTTACACTGAT 480
Qy 478 TATGTAGTACGAGATGGTACCGAGCTCTGAACCTCTTGTGGAGATACCTCAGTATGGT 537
Db 481 TATGTGCAACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 538 TCTTCAGTGCATATATGGCTATTTGGTGTGTTTTCAGAGCTCCGACAGCCAGCCAA 597
Db 541 AAGCTGTGGATGTGTGGGCAATGGTGTGTGTGTAATGAAATGCTATCGGGGAAACC 600
Qy 598 CTGTGGCTGGAATCAGATGTGGACCACTTTATCTGATATCAGAACTAGGAA 657
Db 602 CTGTTTCTGGAGACTCTGATATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 658 TTAATCCAGACATCAATCAATCTTTAAAGTAAACGGGTTTTTCCATGGCATCATATA 717
Db 661 CTAATTCAGACACACAGGAGCTTTTATAAATAATCCTGTGTTGCTGGAGTAAAGT-G 720
Qy 718 CCGAGCCAGNAGACATGGAACCTCTTGAGGAAAGTTCTCAGATGTTTATCTCTGGCT 777
Db 721 CCTGAAATCAAGGAATCAGAACCTCTTTGAAAGACGCTATCCCAAGCTCTCAGAA 780
Qy 778 CTGAATCTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTCTCCCA 837
Db 781 ATAGATTAGCAAGAAATGCTTACATGTTGACCCAGACAAAGGCCCTCTCTGTGTGAG 840
Qy 838 CTCCTGGAGAGCTCCTACTTTGA 860
Db 841 CTCCTACCAATGATTTCTTTCA 863

RESULT 2
US-09-411-628-3
; Sequence 3, Application JS/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOVIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT FILING DATE: 1999-10-01
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3080
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; NAME/KEY: CDS
; LOCATION: (249)...(1949)
US-09-411-628-3

Query Match
Best Local Similarity 59.1%; Pred. No. 2.5e-71;
Matches 510; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

Qy 1 ATGGAAAAGTATGAAAAATAGCTTAAGACTGGAGAGGGCTTATGGGTTGTTATTTCAA 60
Db 249 ATGGAAAAATAGAACTTTGGATTGTTGGAGAGGGAGTTATGAAATGGTATGAAG 308
Qy 61 TGCAGAAACAAACCTCTGGACACAGTACTGTGTAAATAATTTGGGAATCTGAAGAT 120
Db 309 TGTAGAAATAAGATAGTGGAGAGAAATTTGGCCATCAAGAAAGTTCTTAGAAAGTATGAT 368
Qy 121 GATCCTGTTGTTAAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAACAT 180
```

```
Db 369 GACAAAATGGTTAAATAAATTTGCTATCGAGAAATCAAGTTACTAAAGCAACTGAGCAT 428
Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAGGAAAATGCTTTAGTTTTT 240
Db 429 GAAAATTTGGTGAATCTGTTGGAGGTGTGTAATAAATAAATAAATAAATAAATAAATA 488
Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAAACCCAAATGGATTTGCTGAT 300
Db 489 GAATTTGTGACACACGATTTCTGATGACTTGGAACTCTTCCAAATGGACTAGATGAC 548
Qy 301 GAGTGTATCAAAAGCGTATTAATGGCAAAACACTTCAAGCTCTTAAATTTCTGTCATATACAT 360
Db 549 CAAGTAGTTCAAAAGTATTTGTTTTCAGATTAATTAATGGAATTTGATTTTGTCAAGTAC 608
Qy 361 AACTGTATTCACAGAGATATAAACCCTGAAATATTTCTAATTAACCTAAGCAAGGAAATATC 420
Db 609 AATATCATACATAGAGATATAAGCCAGAGAAATATTTGTTCTCCAGTCTGGCGTTGTC 668
Qy 421 AAGATTGTGACTTTCGGGTTTTCACAAAATTTCTG---ATTCCAGAGATGCTTACACCGAT 477
Db 669 AAGTTATCTGATTTTGGATTGTGACGGACACTGSCAGCTCCCGAGAGGTTTACACTGAT 728
Qy 478 TATGTAGTACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
Db 729 TATGTGCAACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 788
Qy 538 TCTTCAGTGCATATATGGGCTATTTGGTGTGTTTTCAGAGACTCTCTGACAGCCAGCCA 597
Db 789 AAGCTGTGGATGTGTGGGCAATTTGGTGTGTTGCTGTAATGAAATGCTCATGGGGAACCC 848
Qy 598 CTGTGGCTGGAATCAGATGTGGACCACTTTTATCTGATTAATCAGAACTAGGAAAA 657
Db 849 CTGTTTCTCGAGACTCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 908
Qy 658 TTAATCCAGACATCAATCAATCTTTAAAGTAAACGGGTTTTTCCATGGCATCATATA 717
Db 909 CTAATTCAGACACACAGGAGCTTTTATAAATAATCCTGTGTTTGTGAGTAAAGTTG 968
Qy 718 CCGAGCCAGNAGACATGGAACCTCTTGAGGAAAGTTCTCAGATGTTTCACTCTGGCT 777
Db 969 CCGTGAATCAAGGAATCAGAACCTCTTGAAGACGCTATCCCAAGCTCTCAGAAAGTTG 1028
Qy 778 CTGAATCTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTCTCCAA 837
Db 1029 ATAGATTTAGCAAGAAATGCTTACATGTTGACCCAGACAAAGGCCCTCTCTGTGTGAG 1088
Qy 838 CTCCTGGAGAGCTCCTACTTTGA 860
Db 1089 CTCCTACCAATGATTTCTTTCA 1111

RESULT 3
US-08-874-347-9
; Sequence 9, Application JS/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```



```
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,347
FILING DATE: 13-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...900
OTHER INFORMATION:
US-08-874-347-9

Query Match      18.2%; Score 171.8; DB 2; Length 903;
Best Local Similarity 52.3%; Pred. No. 5e-39;
Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;

QY 1 ATGGAAAGTATGAAATTAAGTCTGAGAGGAGGCTTTATGGGTGTGTTATCAAA 60
DB 1 ATGGAGCAATATCAGAGGTTCAGAGAGATGAGAGGAGGAGGCTTTATGAGGAT 60
QY 61 TCGAGAAACAAACCTCGAGCAAGTAGTAGCTTTAAATAAATTGTGGAATCTGAAGAT 120
DB 61 GCAGAGGATCTTGAAGTGGTACAAATTGTAGCTTTAAGAAATCCGATTAGAGACGAGAA 120
QY 121 GATCCTGTGTTAAGAAATAGCCTAAGAGAAATACGTAAGTTTGAAGCAATTAACAT 180
DB 121 GATGAGGAGGAGTCTTAGTAGACAGCAATTCGTGAGATATCACTTTTGAAGAGATGCACAAT 180
QY 181 CCAATATCTGTCAACCTCATCGAGGTGTCAGGAGAGAAAGGAAATGCATTAGTTT 240
DB 181 GATAATGTTGTGAAGACTTTTGAATATAATTCATCAAGAGTCAAGTTTATATCTGTTT 240
QY 241 GAATACCTGTGATCATAACACTTTTAAATGAGCTGGAAGAAACCCAAATG-----GAGTT 294
DB 241 GAATTTCTTGAATCTTGAATTTAAAGGATATATGATAGTATTTCCAAAGGACATGCTT 300
QY 295 GGTGATGAGTATCAAGAGCTATATGGGAAACACTTCAAGCTTATTTCTGTGAT 354
DB 301 GTGAGAGAAATGATTAAAGGTTTATGTACAACTGTTATCAGGTGTAAATATTTGTCAT 360
QY 355 ATACATAACTGTATTCACAGAGATATAAAGCTGAAATATTTCTAACTAAGCAAGGA 414
DB 361 TCTCATGCTATTCTTCATGCTGACTTGAACCAACAAATCTTTATAGATCGAGAGGA 420
QY 415 ATAAATCAAGATTGTGACTTCGGGTTTGA-----CAATTTCTGATTCCAGGAGATGCCTAC 471
DB 421 AATCTTAAATATGACAGATTTCGGGCTTGCAGGGGCTTTGCTGTTCCATTCGGTGGTAT 480
QY 472 ACCGATTATGTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATACCTAC 531
DB 481 ACTCATGAAGTTGTTACACTTGGTATCGTCTCAGAGATGTTCTTTAGGTGGTCGACAA 540
QY 532 TATGGTCTTCTAGTTCGATATGGGCTATTTGGTTGTTTGTGTTTTCAGAGCTCTGACAGGC 591
DB 541 TATGCAACAGCGCTTCATATATGGAGCATTTGGATGTATTTTTCAGAGAAATGGCTACAAA 600
QY 592 CAGCCACTGTGGCTGGAAATACAGATGAGGACCAACTTTATCTGATTAATCAGAACACTA 651

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,347
FILING DATE: 13-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...900
OTHER INFORMATION:
US-09-671-050-11.rni

Query Match      18.2%; Score 171.8; DB 3; Length 903;
Best Local Similarity 52.3%; Pred. No. 5e-39;
Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;

QY 601 AAGCCATTATTTCCAGGTGATTTCTGAAATTCATGAAATATTAGAAATTATTAGAAATTATTA 660
QY 652 GGAAATTAATCCCAAGACATCAATCAATCTTTAAAGTAAAGGTTTTCCTCATGGGATC 711
DB 661 GGGACTCCAGATGAAATTTCTTGGCTGGTATTACATCTTTATCCGGATTTTAAAGGCACT 720
QY 712 AGTATACCTGAGCCAGAGACATCGAAACTCTTCAGGAAAGATTCTCAGATGTTTCATCCT 771
DB 721 TTTCCAAATGGTCACCA-----AAAAATCTTGGAGATTAATACAGAACTTGATAGT 774
QY 772 GTGGCTCTGAACTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACTGT 831
DB 775 GATGATATAGATTTATTACAGAAATGCTTAGATATTATCTCTGCTGAACGTTATTAGGCT 834
QY 832 TCCCAACTCTCGGAGAGCTCTCTACTTTGATTTCTTT 866
DB 835 AAAAAAGCTCTCGATCATCTCTTATTTTGTATGATTT 869

RESULT 4
US-09-671-050-11.rni
Sequence 9, Application US/09093522
Patent No. 6015700
GENERAL INFORMATION:
APPLICANT: Limper, Andrew H.
APPLICANT: Leof, Edward B.
APPLICANT: Thomas, Charles F.
APPLICANT: Gustafson, Michael P.
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
TITLE OF INVENTION: CAELINI
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSES: Fish & Richardson P.C., P.A.
STREET: 60 South Sixth Street, Suite 3300
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/093,522
FILING DATE: 08-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/874,347
FILING DATE: 13-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...900
OTHER INFORMATION:
US-09-093-522-9
```

Best Local Similarity 52.3%; Pred. No. 5e-39;	
Matches 458; Conservative	0; Mismatches 402; Indels 15; Gaps 3;
QY	1 ATSGAAAAAGTATGAARAAATTAGCTAAGACCTGGAGAGGGTCCTTATGGGTCCTGATTCGAA 50
Db	1 ATGGAGCAATATCAGAGGTTAGAGAGATTTGGAGAGGAACCTTATGGAGTTGTTTATAAA 60
QY	61 TCGAGAAACAAACCTCTGGACAAGTASTAGCTTTAAAAAATTTGTGSAATCTGAAAGT 120
Db	61 GCAAAGGATCTTGAAGTGGTACAATTGTAGCTTTAAGAAAATCCGATTAGAAGCAGAA 120
QY	121 GATCTGTGTTTAAAGAAATAGCACTTAAGAGAAATAGTATGTGGAGCAATTAACAT 180
Db	121 GATGAGGGAGTTCTCTAGTACAGCAATTCGTGAGATACTCTTTGAAGAGAGTGCACAACT 180
QY	181 CCAAAATCTTTGAAACCTCATCGAGGTGTTTCAGGAGAAAAAGAAAATGCATTAGTATTTT 240
Db	181 GATAATGTTTGAAGACTTTTGAATATAATTATCATCAGAGTCAGCTTTATATCTTGTTTT 240
QY	241 GAATACTGTGATCATACACTTTTAAATAGAGCTGAAAGAAACCCAAATG-----GAGTT 294
Db	241 GAATTTCTGTATCTTGATTTGATTTAAAAAGTATATGAATAGTATTCCAAAGACATGATGCTT 300
QY	295 GCTGATGGAGTATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTGAT 354
Db	301 GGTGCAGAAATGNTTAAAGTTTATGTCACACTTGTATCAGGTCGTAATATTGTCAT 360
QY	355 ATACATAACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGA 414
Db	361 TCTCATCGTATTTCTTCATCGTGACTTGAACCCACAAAATCTCTTATAGATCGAAGGA 420
QY	415 ATAAATCAAGATTTGTGACTTCGGGTTTGCA---CAAATTCGATTCGAGAGATGCTTAC 471
Db	421 AATCTTAAATTAGCAGATTTTGGGCTTGCAAGGCGGTTTGGTGTCCATTTGGTGGTTAT 480
QY	472 ACGGATTTATGTAGCTACGAGATGGTACCGAGCTCTGAACTCTCTGTGGAGATACTCAG 531
Db	481 ACTCATGAAGTTGTTACACTTTGGTATCGTGCTCCAGAAGTCTTTTAGTGGTTCGACAA 540
QY	532 TATGGTCTTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTTCGAGAGCTCTGCACAGC 591
Db	541 TATGCAACAGCCTTGATATATGGAGCATTTGGATGTATTTTTCAGAAAATGCTCACAAA 600
QY	592 CAGCCACTGTGGCCTGGAAAACTCAGATGTGACCAACTTTATCTGATATATCAGAACACTA 651
Db	601 AAGCCATTATTTCCAGTGTGATCTGAAATGTATGAAATATTTAGATATTTAGATATTA 660
QY	652 GGAATAATTATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTCCCATGGCAT 711
Db	661 GGGACTCCAGATGAAAATTTCTTGGCCTGGTATTACATCTTATCCGGATTTTAAGGCACT 720
QY	712 AGTATACCTGAGCCAGAGACATGMAACTCTTCAGGAAAAGTTCTCAGATGTTTATCCT 771
Db	721 TTTCAAAATGTCACCA-----AAAAATCTTGGAGAAATTAAATACAGACTTGATAGT 774
QY	772 GTGGCTCTGAACTTCATGAAGGGGTGTCGTAAGATGAATCCAGATGACAGATTAACCTGT 831
Db	775 GATGGAATAGATTTTATTACAGAAATGCTTAGATATTATCTCTGCTGAACGATATACCGCT 834
QY	832 TCCCAACTCTCTGGAGAGTCTCTACTTTGATTCCTTT 866
Db	835 AAAAAAGTCTCTGATCATCCCTTAATTTTATGATGATTT 869


```
FEATURE:
; NAME/KEY: CDS
; LOCATION: (17..(1632)
US-09-417-197-114

Query Match      17.3%; Score 163.4; DB 4; Length 1635;
Best Local Similarity 55.1%; Pred. No. 1.6e-36;
Matches 364; Conservative 0; Mismatches 291; Indels 6; Gaps 2;

QY 1 ATGGAAGATGTAAGAAATTAGCTAGAGCTGGAGAGGGTCTTATGGGTGTGTTCAAA 60
DB 739 ATGGAGAACTTCCAAAGGTGGAAGATCGGAGAGGCGCTACGGAGTTGTGTACAAA 798
QY 61 TCAGAGAAACAAACCTCTGGCAAGCTAGTAGCTGTAAAGAAATTTGTGGAATCTGAAGAT 120
DB 799 GCCAGAAACAAGTTGACGGGAGAGGTGTGGCGCTTAAGAAATCCCGCTGGACACTGAG 858
QY 121 GATCTGTGTTAAGAAATAGCACTAGAGAAATACGTATGTTGAAGCAATTAAGACAT 180
DB 859 ACTGAGGGTGTGCCAGTACTGCCATCCGAGAGATCTCTCTCTTAAGGAGCTTAACCAT 919
QY 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCAATTTAGTTT 242
DB 919 CTTAATATTTCAAGTCTGATGTGATCTCACAGAAATATACTCTACTCTGTTTT 978
QY 241 GAATAGTGTATACACTTTTAAATGAGTGA---AAGAAACCCAAATGSGAGTTGCT 297
DB 979 GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCTCTGCTCTCACTGGCATTCCT 1038
QY 298 GATGAGTGTATCAAGAGCTATTATGGGAAACACTTCAGCTCTTAA---TTCTGTCTATATA 357
DB 1039 CTTCCCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGSCCTAGCTTTCTGCCATCTCT 1098
QY 358 CATAACTGTATTACAGAGATATAAACCCTGAAATATTTTAAATCACTAAGCAAGGAATA 417
DB 1099 CATGGGCTCTCCACGAGACCTTAAACCTCAGATCTGCTTATTAAACACAGAGGGGCC 1158
QY 418 ATCAAGATTTGTAGCTTCGGTTTGCACAAATCTGA---TTCAGGAGATGCTCTACACC 474
DB 1159 ATCAAGCTAGCAGACTTTGAGCTAGCAGAGCTTTGAGAGTCTCTGTCGTAATACACC 1219
QY 475 GATTATGTAGCTACGAGATGTTACGAGCTCTGGAACCTCTTGTGGAGATACCTAGTAT 534
DB 1219 CATGAGTGTGACCTGTGTTACCGAGCTCTGGAATCTCTCTGGCTCGAAATATATAT 1278
QY 535 GGTCTTTCAGTCTGATATATGGCTATTTGTTGTTTGTGAGAGCTCTCTGACAGGGCAG 594
DB 1279 TCCACAGCTGTGGACATCTGGAGCCTGGGCTGCATCTTTGCTGAGATGGTGACTCGCGG 1338
QY 595 CCACTGTGGCCTGGAAATACAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGA 654
DB 1339 GCCTGTCTCCCTGGAGATTTGAGATTTGACAGCTCTTCCGAGTCTTTTCGAGCTCTGGG 1398
QY 655 A 655
DB 1399 A 1399

RESULT 10
US-09-220-132-3
; Sequence 3, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shivjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220.132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24

; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-3

Query Match      16.5%; Score 155.6; DB 4; Length 1050;
Best Local Similarity 54.4%; Pred. No. 2e-34;
Matches 361; Conservative 0; Mismatches 294; Indels 9; Gaps 2;

QY 1 ATGGAAGATGTAAGAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTGTGTTCAAA 60
DB 127 ATGGAAGATTTATCCAAATAGAGAAATTTGGAAGGTACCTATGGAGTTGTGTAAAG 186
QY 61 TCAGAAACAAACCTCTGGACAGTAGTAGCTCTTAAGAAATTTTGGATCTGAAGAT 120
DB 187 GGTAGACACAAACTTACAGGTCAAGTGTGATGCCATGAAAGAAATTCAGACTAAGAGTGA 246
QY 121 GATCCTGTGTTTAAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAAGACAT 180
DB 247 GAGGAAGGGGTCTCTAGTACTGCAATTCGGGAATTTCTCTATTAAAGGAACCTTCGTCAT 306
QY 181 CCAAACTCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCAATTTAGTTT 246
DB 307 CCAAAATATAGTCAGTCTTTCAGGATGTGCTTATGAGGATTCAGGTTTATCTCATCTTT 366
QY 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGG-----AGTT 294
DB 367 GAGTTTCTTCCATGGATCTGAAGAAATCTTGGATTTCTATCCCTCTGGTCACTACATG 426
QY 295 GCTGATGAGTGTACAAAGCGTATTATGCGAAACACTTTCAGCTCTTAA---TTCTGTCTCAT 354
DB 427 GATTTCTTCACTTGTTAAGAGTATTATTATACCAATCTCTACAGGGGATTTGTTTGTGAC 486
QY 355 ATACATACTGTATTACAGAGATATAAACCCTGAAATATTTCTAATAACTAAGCAAGGA 414
DB 487 TTAGAAGAGTTCTTACAGAGACTTAAACCTCAAAATCTCTTGATTGATGACAAAGGA 546
QY 415 ATAATCAAGATTTGTGACTCTGGGTTTGCACAAATTTCTG---ATTCCAGGAGATGCCCTAC 471
DB 547 ACAATTAACCTGCTGATTTTGGCCTTCCAGAGCTTTTGGAAATACCTATCAGAGTATAT 606
QY 472 ACGGATATGATGATACAGAGATGTTACCGAGCTCTGAACTCTTGTGGAGATACCTCAG 531
DB 607 ACACATGAGGTAGTAACACTCTGTTACAGATCTCCAGAAAGTATTGCTGGGTGAGCTCGT 666
QY 532 TATGGTCTTTCAGTCTGATATATGGGCTATTGGTTGTTGTTTTGAGAGCTCTCTGACAGGC 591
DB 667 TACTCACTCCAGTTGACATTTGGAGTATAGCACCATTATTTGCTGAACCTAGCACTAAG 726
QY 592 CAGCACTGTGGCCTGGAAATTCAGATGTGACCAACTTTTATCTGTGATAATCAGAACTA 651
DB 727 AAACCACTTTTCCATGGGATTCAGAAATTTGATCAACTCTTCAGGATTTTCAGAGCTTTG 786
QY 652 GGAA 655
DB 787 GGCA 790
```

```
RESULT 11
US-09-620-312D-313
; Sequence 313, Application US/09620312D
; Patent No. 6596662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenchua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
```

```

; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aihong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillingshast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6563662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PL_FL_genes Version 1.0
; SEQ ID NO 313
; LENGTH: 1825
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (234)...(1235)
; US-09-620-312D-313

Query Match      14.5%; Score 136.6; DB 4; Length 1825;
Best Local Similarity 52.5%; Pred. No. 6.1e-29;
Matches 348; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

Qy      1 ATGGAAGAATGATGAAATAGCTAAGCTGAGAGGGCTTATGGGTGTGATTCAAA 63
Db      1 ATGGAAGAATGATGAAATAGCTAAGCTGAGAGGGCTTATGGGTGTGATTCAAA 63

Qy      318 ATGCATATGTCAGAGAGGTAGAGAGATCGAGAGGGCACCTATGGGTGTGATCAAG 377
Db      318 ATGCATATGTCAGAGAGGTAGAGAGATCGAGAGGGCACCTATGGGTGTGATCAAG 377

Qy      61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTTAAATAATTTGGAAATCTGAAGAT 120
Db      61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTTAAATAATTTGGAAATCTGAAGAT 120

Qy      378 GCCAAGAACAGGGAGACAGGGGACCTGCTGGCCCTGAAGAAGATCAGACTGGATTGGAG 437
Db      378 GCCAAGAACAGGGAGACAGGGGACCTGCTGGCCCTGAAGAAGATCAGACTGGATTGGAG 437

Qy      121 GATCCTCTGTTGAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAGACAT 180
Db      121 GATCCTCTGTTGAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAGACAT 180

Qy      438 ATGAGGGGGTCCCAAGCACTGCATCAGGGAGATCTCGCTGCTCAAGNACTGAAGCAC 497
Db      438 ATGAGGGGGTCCCAAGCACTGCATCAGGGAGATCTCGCTGCTCAAGNACTGAAGCAC 497

Qy      181 CCAAAATTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAGGAAATGCAATTTAGTTTTT 243
Db      181 CCAAAATTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAGGAAATGCAATTTAGTTTTT 243

Qy      498 CCCAACATCGTCCGACTGCTGGAGTGTGTCACAAACGAGAGAGAGCTCTATCTGGTGT 557
Db      498 CCCAACATCGTCCGACTGCTGGAGTGTGTCACAAACGAGAGAGAGCTCTATCTGGTGT 557

Qy      241 GAATGCTGTGATCATACACTTTTAATGAGCTGGAAGAAACCCAAATGGAG--TTGCT 297
Db      241 GAATGCTGTGATCATACACTTTTAATGAGCTGGAAGAAACCCAAATGGAG--TTGCT 297

Qy      558 GAGTTCCTCAGCAGGACCTGAGAGAGTACATGACCTCCACCCAGGCTCAGAGCTCCCC 617
Db      558 GAGTTCCTCAGCAGGACCTGAGAGAGTACATGACCTCCACCCAGGCTCAGAGCTCCCC 617

Qy      298 GATGGAGTATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGCATATA 357
Db      298 GATGGAGTATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGCATATA 357

Qy      618 CTGCACCTCATCAAGAGTACCTCTTCAGCTGCTGCGAGGGGTGAGTTCTCCCACTCA 677
Db      618 CTGCACCTCATCAAGAGTACCTCTTCAGCTGCTGCGAGGGGTGAGTTCTCCCACTCA 677

Qy      358 CATTAAGTGTATTCACAGAGATATAAAACCTGGAATATTTATTAAGTAAGCAAGGAATA 417
Db      358 CATTAAGTGTATTCACAGAGATATAAAACCTGGAATATTTATTAAGTAAGCAAGGAATA 417

Qy      678 CATCGGTATCCACCGAGACCTGAAGCCCAAGAACTGCTCATCAATGAGTTGGGTGCC 737
Db      678 CATCGGTATCCACCGAGACCTGAAGCCCAAGAACTGCTCATCAATGAGTTGGGTGCC 737

Qy      418 ATCAAGATTGTGACTTCGGGTTCGACAAATCT--GATTCAGAGAGATGCTACACC 474
Db      418 ATCAAGATTGTGACTTCGGGTTCGACAAATCT--GATTCAGAGAGATGCTACACC 474

Qy      738 ATCAAGCTGGTGAATTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 797
Db      738 ATCAAGCTGGTGAATTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 797

Qy      475 GATTATGTAGTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGAGATATCTAGTAT 534
Db      475 GATTATGTAGTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGAGATATCTAGTAT 534

Qy      798 CATGAGGTGGTGAACACTGTGTATCGCGCCCCGAGATTCTCTTGGGACCAAGTTCTAT 857
Db      798 CATGAGGTGGTGAACACTGTGTATCGCGCCCCGAGATTCTCTTGGGACCAAGTTCTAT 857

Qy      535 GGTTCTTCAGTCGATATATGGGCTATTGGTTGTTTTTSCAGAGCTCTTCAGAGGCCAG 594
Db      535 GGTTCTTCAGTCGATATATGGGCTATTGGTTGTTTTTSCAGAGCTCTTCAGAGGCCAG 594

Qy      858 ACCACAGCTGTGGATATCTGGAGCAATGGTTGATCTTTTCAGAGATGGTGAATCGAAAA 917
Db      858 ACCACAGCTGTGGATATCTGGAGCAATGGTTGATCTTTTCAGAGATGGTGAATCGAAAA 917

; 595 CCACGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGA 654
; 918 GCCGTGTTCTGGTGACTCTGAGATTGACCAAGCTCTTTCTGATCTTTCTGATCTGGGG 977
; 655 AAA 657
; 978 ACA 980

RESULT 12
US-08-154-915-1
; Sequence 1, Application US/08154915
; Patent No. 5618669
; GENERAL INFORMATION:
; APPLICANT: Beach, David
; APPLICANT: Xiong, Yue
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses
; TITLE OF INVENTION: Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/154,915
; FILING DATE: 19-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1993
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
; US-08-154-915-1

Query Match      14.2%; Score 134.6; DB 1; Length 1089;
Best Local Similarity 52.3%; Pred. No. 1.8e-28;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

Qy      1 ATGGAAGAATGATGAAATAGCTAAGAGAGGGTCTTATGGGTGTGATTCAAA 60
Db      1 ATGGAAGAATGATGAAATAGCTAAGAGAGGGTCTTATGGGTGTGATTCAAA 60
```

13 ATGCAGAAATACGAGAAATCGAAAGATTGGGAAGGCACCTACGGAACCTGTGTTCAAG 72
61 TGCAGAAACAAACCTCTGGACAGTACTAGCTGTAAAGAAATTTGGGATCTGAGAT 120
73 GCCAAACACCGGAGACTCATGAGATCTGTGGCTTAAACCGGTGAGCTGGATGAGCAT 132
121 GATCTGTGTTTAAAGAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACACAT 180
133 GATGAGGTGTGCCGAGTTCGCCCTCCGGGAGATCTGCCCTACTCAAGGAGCTGAACAC 192
181 CCAATCTGTGAACCTCATCGAGTGTTCAGGAGAAACGAAATGCAATTTAGTTT 240
193 AAGAACATCTGTCAGGCTTCATGAGCTCTGCACAGCGACAAAGAGCTGACTTTCGTTT 252
241 GAATCTGTGATCATACACTTTAAATGAGCTGGGAAGAAACCCCAATGGAGTTGCTGAT 300
253 GAATTCGTGACCAAGGACCTGAAGAAGTATTTGACAGTTCCAAATGGTGACCTGCATCT 312
301 GGAGTGATCAAGAGCTTATTGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACAT 360
313 GAGATGTAAAGTCACTGTGATACCGCCCAAGATGCTCTTTGGGCGCAAGCTGTACTCC 372
361 AACTGTATTCAGAGATATAAAGCTGAAATATCTTAATTAAGCAAGGAATCAATC 420
373 AATGTGCTACACAGGACCTGAAGCCCGCAGAACCTGCTAATACAGGAATGGGAGCTG 432
421 AAGATTGTGACTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 477
433 GAGTGATCAAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 492
478 TATGTAGCTACGAGATGATACCGAGCTCCTGAACCTCTTAAATTTCTGTCTATATACAT 537
493 GAGTGATCAAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 552
538 TCTTCAGTTCGATATATGGGCTATTTGGTGTGTTTTCAGAGCT---CCTGACAGGCGAG 594
553 ACGTCCATCGACATGTGTCAGCCGCTGCTATCTTTCAGAGCTGGGCGGCTGCTGGG 612
595 CCAGTGTGGCTGGAAATCAGATGTGACCAACTTTATCTGATATCAAGCAACACTAGGA 654
613 CCTCTTTTCCCGCAATGATGTGATGACCACTTGAAGAGGATCTTCCGACTGCTGGG 672

RESULT 13

US-08-464-517-37

Sequence 37, Application US/08464517

Patent No. 5869640

GENERAL INFORMATION:

APPLICANT: BEACH, David H.

TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSER: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/464,517

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/963,308

FILING DATE: 16-OCT-1992
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: M11-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7403
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13...888
US-08-464-517-37

Query Match: 14.2%; Score 134.6; DB 2; Length 1089;

Best Local Similarity 52.3%; Pred. No. 1.8e-28;

Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 120
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 120
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 180
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 180
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 240
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 240
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 300
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 300
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 360
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 360
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 420
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 420
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 477
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 477
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 537
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 537
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 594
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 594
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 654
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 654
QY : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 672
DB : ATGCAAAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 672

QY 655 A 655
Db 673 A 673

RESULT 14
US-08-246-361A-37
; Sequence 37, Application US/08246361A
; Patent No. 5998582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THEREO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246.361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: M11-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-5940
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-246-361A-37

Query Match 14.2%; Score 134.6; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 1.8e-28;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY 1 ATGGAAGAGTATGAAAGAAATAGCTAAGACTGGAGAGGGCTTATGGGTTGTATTCAAA 60
Db 13 ATGCAGAAATACAGAACTGGAAAGAGATTGGAGAGGCACTACGAGACTGTGTTCAAG 72

QY 61 TCAGAGAAACAACTCTGGAGCACTAGTACTGTGTTAAAGAAATTTGGGAATCTGAAGAT 120
Db 73 GCCAAAGCCGGAGACTCATGAGATCTGGCTCTAAAGCCGGTGGAGCTGGATGACCAT 132

QY 121 GATCTGTTGTTAAGAAATACCTAGAGAAATACGATATGTTGAAGCAATTAACAT 180
Db 133 GATGAGGGTGTCCCGAGTTCCCGCTCCCGGAGATCTGCTACTCAAGGAGCTGAAGCAC 192

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGAGAGAAAGGAAATGCATTTAGTTTTT 240
Db 193 AAGAAATCTGTGAGGCTTCATGACGCTCTGCACAGCGACAAGAGCTGACTTTGGTTTTT 252

QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATAGAGTTGCTGAT 300
Db 253 GAATCTGTGACAGGACCTGGAAGAGTATTTTGCAGATTGCAATGGTGACCTCGATCCT 312

QY 301 GGACTGATCAAAAGCGTATTATGGCAACACTTCAGCTCTTAATTTCTGTCTATACAT 360
Db 313 GAGATTGTAAGTCATTCCTCTTCCAGCTACTAAAGGGCTGGGATTCGTCTATGCCGC 372

QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGCAATAATC 420
Db 373 AATGTGCTACACAGGACCTGAAAGCCAGAACCTGCTAATAAACAGGAATGGGAGCTG 432

QY 421 AAGATTTGTGACTTCGGGTTTGACAAA---TTCTGATTCAGGAGATGCCTACACCGAT 477
Db 433 AAATTGGCTGATTTGGCTGGCTGGAGGCTTTGGGATTTCCCGTCCGCTGTACTCAGCT 492

QY 478 TATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTTCTGTGGGAGATCTCAGTATGCT 537
Db 493 GAGTGTGACACTGTGCTACCCACCGAGTCTCTTTGGGGCCAAAGCTGTACTCC 552

QY 538 TCTTCAGTGCATATATGGGCTATTTGTTGTTTGTGTCAGAGCT---CCTGACAGCCAG 594
Db 553 ACGTCCATCGACATGTGCTCAGCGGCTGCTCTTTGTCAGAGCTGGCCAAATGCTGGCGG 612

QY 595 CCAGTGTGGCTGGAATATCAGATGGGACCACTTTATCTGATATATCAGAACACTAGGA 654
Db 613 CCTCTTTTCCCGCAATGATGCGATGACCACTTGAAGAGGATCTTCGACTGCTGGGG 672

QY 655 A 655
Db 673 A 673

RESULT 15
US-08-463-772-37
; Sequence 37, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THEREO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: M11-004C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 1089 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 13..889

US-08-463-772-37

Query Match 14.2%; Score 134.6; DB 3; Length 1089;

Best Local Similarity 52.3%; Pred. No. 1.8e-28;

Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

```
QY 1 ATGGAAAGTATGAAATAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 13 ATGCAGAAATACGAGAACTGAAAGAATGGGAAGGCACCTACGGAACCTGTGTTCAAG 72
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TGCAGAACCAAAACCTCTGGACAGCTAGCTGTTAATAAATTTTGGATCTGAGAT 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 73 GCCAAACCCGGAGACTCATGAGATCGTGGCTCTAAACACGGGTGAGCTGGATGACGAT 132
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GATCTGTGTTTGAAGAAATAGCACTAAGAGAAATACGTATGTTGAACCAATTAACACAT 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 133 GATGAGGGTGTCCGAGTTCGCCCTCCGGGAGATCTGCCCTACTCAAGGAGCTGAAGCAC 192
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTGAGGAGAAAGAAATGCAATTTAGTTTTT 240
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 193 AAGACATCGTCAGSCTTCATGACGCTCTGCACACGCGACAGAAAGCTGACTTTGGTTT 252
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 GAATACGTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT 300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 253 GAATCTGTGACCAAGGACTGTGAGAGTATTTTGACAGTTGCAATGGTGACCTGATCCT 312
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 GGAGTGATCAAAAGGGTATTTATGGGAAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 313 GAGATTGTAAGTCAATTCCTCTTCCAGCTACTTAAAGGGCTGGGATTCGTTCATAGCCGC 372
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 AACTGTATTACAGAGATATAAAGCTGAATATTTCTAATAACTTAAGCAAGGAATATC 420
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 373 AATGCTACAGGACCTGAAGCCCGAAGACCTGCTTAATAACAGGAATGGGAGCTG 432
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 AAGATTGTGACTTCGGGTTTCACAAA---TTCTGATTCAGAGAGATGCCCTACACCGAT 477
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 433 AAATTGGCTGATTTGGCCCTGGCTCGAGSCTTTGGGATTCGCCGCTGTTACTCAGCT 492
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 TATGTAGCTACGAGATGTTACCGAGCTCTGAACCTTTGTTGGGAGATACCTCAGTATGGT 537
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 GAGGTGGTCACACTGTGGTACCGCCACCGGATGTCCTCTTTGGGGCCAAAGCTGTACTCC 552
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 TCTTCAGTCGATATATGGGCTATTTGGTTGTGTTTTCAGAGCT---CCTGACAGGCCAG 594
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 553 ACCTCCATCGACATGTGGTCAGCCGGCTGCATCTTTGACAGCTGGCCAAATGCTGGGGGG 612
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 CCACTGTGGCCTGGAAATACAGATGTGACCAACTTTTACTTGATTAATCAGAACACTAGGA 654
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 613 CTTCTTTTCCCGCAATGATGTGATGACCAAGTTGAAGAGGATCTTCCGACTGTGGGG 672
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 655 A 655
Db 673 A 673
```

Search completed: November 15, 2003, 22:31:38

Job time : 70.0921 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw mode?

Run on: November 15, 2003, 16:06:46 : Search time 2379.15 seconds
(without alignments)
9653.738 Million cell updates/sec

Title: US-09-671-050-11

Perfect score: 945

Sequence: 1 atggaagatgataaaatt.....aggtaattccgtcaaaagt 945

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum Match 100%

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estm.*

5: em_estov.*

6: em_estp.*

7: em_estro.*

8: em_estc.*

9: gb_estl.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfun.*

16: em_estom.*

17: em_gss_rum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gssl.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
C 1	480.6	50.9	695	14	CB169554	CB169554 RUC603000
2	405	42.9	670	14	BY733578	BY733578 BY733578
3	401.8	42.5	1691	11	AK016781	AK016781 Mus muscu
4	390.6	41.3	825	14	CB315094	CB315094 AGENCOURT

C	5	361	38.2	536	13	BU686325	BU686325 UI-CF-DUI
6	360.2	38.1	689	14	BY752719	BY752719 BY752739	
7	333.6	35.3	500	9	AI385966	AI385966 m134h09.y	
8	319.4	33.8	824	13	BU221831	BU221831 603750354	
9	316.8	33.5	938	9	AA061797	AA061797 m134h09.r	
10	306.8	32.5	598	9	AV986182	AV986182 AV986182	
11	289.2	30.6	597	12	BP016186	BP016186 BP016186	
12	269.8	28.6	565	14	CB400506	CB400506 OSTF177FL	
13	258.6	27.4	731	9	AW106692	AW106692 um32d03.y	
14	255	27.0	645	14	CA334835	CA334835 NISC-1t01	
15	255	27.0	718	9	AA286088	AA286088 vc33b05.r	
16	250.6	26.5	2572	11	AK030598	AK030598 Mus muscu	
17	249	26.3	508	12	BI511582	BI511582 BX160006A	
18	247.8	26.2	512	13	EX304025	EX304025 BX304025	
19	239	25.3	485	9	AV960213	AV960213 AV960213	
20	231.2	24.5	812	13	BU377342	BU377342 603812152	
21	227.2	24.0	1064	13	BU120363	BU120363 603142234	
22	226	23.9	774	13	BW110560	BW110560 BW110560	
23	224.2	23.7	2005	11	AK090045	AK090045 Mus muscu	
24	224.2	23.7	2769	11	AK045036	AK045036 Mus muscu	
25	223.8	23.7	688	13	BW241505	BW241505 BW241505	
26	222.6	23.6	1893	11	AK050990	AK050990 Mus muscu	
27	221	23.4	2388	11	AK046394	AK046394 Mus muscu	
28	221	23.4	2665	11	AK045356	AK045356 Mus muscu	
29	218.6	23.1	750	14	CB955302	CB955302 AGENCOURT	
30	216.2	22.9	432	14	CB757579	CB757579 AWMNNUC:N	
31	213.4	22.6	570	12	BJ000534	BJ000534 BJ000534	
32	211.6	22.4	2013	11	AK078804	AK078804 Mus muscu	
33	210.8	22.3	627	14	BY716459	BY716459 BY716459	
34	209.8	22.2	579	9	AW233105	AW233105 f128b10.y	
35	209.4	22.2	380	13	BY075395	BY075395 BY075395	
36	208.4	22.1	2604	11	AK052380	AK052380 Mus muscu	
37	205.4	21.2	295	9	AA626859	AA626859 zu89f10.s	
38	197.2	20.9	471	14	CB732963	CB732963 AWMNNUC:N	
39	196.9	20.8	851	13	BU463954	BU463954 60369666	
40	192.4	20.4	1535	11	BC010966	BC010966 Homo sapi	
41	188.6	20.0	659	13	BW280090	BW280090 BW280090	
42	187	19.8	673	13	BW036816	BW036816 BW036816	
43	186.6	19.7	727	9	AI508835	AI508835 vc33b05.y	
44	186.2	19.7	666	14	BY733092	BY733092 BY733092	
45	182	19.3	320	13	BY121897	BY121897 BY121897	

ALIGNMENTS

RESULT 1
CB169554/c
LOCUS RUC603000752.R1 CSEQFXN20 testes Bos taurus cDNA, mRNA sequence.
DEFINITION 695 bp mRNA linear EST 30-JAN-2003
ACCESSION CB169554
VERSION CB169554.1 GI:28155681
KEYWORDS EST..
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE
i. (bases 1 to 695)
Adelson,D.L. and Gill,C.A.
Bovine ESTs (Adelson and Gill)
Unpublished
JOURNAL
COMMENT
Contact: David L. Adelson
Animal Breeding and Genetics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
USA
Tel: 9798452616
Fax: 9798456970
Email: david.adelson@amu.edu.
Location/Qualifiers
1..695
/organism="Bos taurus"

```

/mol_type="mrna"
/db_xref="taxon:9913"
/tissue_type="testes"
/clone_id="CSEQXN20 testes"
/note="Organ: testes; Vector: pBluescript SK+; Site: 1:
NotI; Site 2: EcoRI; sequence 5' of the insert
(5'-NNN...NNNinsert)
CGCAATTGGAGCTCCACCGGCTGGCGGCGGCTCGAG. Sequence 3' of
the inserts (AAGAAATCGATATCAAGCTTATGATACCGCTCGACCTCGAG
normalized Rd 1 library, sequenced 3' with M13R primer."
BASE COUNT      169 a 166 c 133 g 227 t
ORIGIN
Query Match      50.9%; Score 480.6; DB 14; Length 695;
Best Local Similarity 92.8%; Pred. No. 5.9e-96;
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1 ATGGAAGATGAGAAAATTAGCTAAGCTGGAGAGAGGCTTTATGGGTTGTTATCAA 60
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
543 ATGGAGAGATGAGAAAATTAGCTAAGCTGGAGAGAGGCTTTATGGGTTGTTATCAA 484
QY 61 TGCAGAAACAAACCTCGGACAAAGTAGCTCTTAAAAAATTGTGGAAATCGAAGAT 120
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
483 TGCAGAAACAAACCTCGGACAAAGTAGCTCTTAAAAAATTGTGGAAATCGAAGAT 424
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GATCCTGTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 190
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
423 GATCCTGTGTTAAGAAAATAGCTCTAAGAGAAATACGCATGTTAAAGCAATTAAACAT 364
QY 181 CCAATCTGTGACCTCATCGAGTGTTCAGGAGAAAGGAAATGCAATTAGTTT 240
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 CCAATCTGTGACCTCATCTGAGTGTTCAGGAGAAAGGAAATGCAATTAGTTT 304
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCAATGAGTTGCTGAT 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 GASTACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCAATGAGTTGCTGAT 244
QY 301 GGAGTGATCAAAAGGTAATATGCAACACACTTCAAGCTCTTAATTTCTGTCTATACAT 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 GGAGTGATCAAAAGTGTGTTATGCAACACACTTCAAGCTCTTAATTTCTGTCTATACAT 184
QY 361 AACTGTATTTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGAAATATC 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 AATGTATTTCACAGAGATGTAATACCTGAAATATTTCTAATACTAAGCAAGAAATATC 124
QY 421 AAGATTGTGACTTGGGTTTGCACAAATTCGATTCAGAGATGCTACACCGATTAT 490
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 AAGATTGTGACTTGGGTTTGCACAAATTCGATTCAGAGATGCTACACCGATTAT 64
QY 481 GTAGTACAGATGTCACGAGCTCCTGAATCTCTGTGGAGATCTCAGTATGTTCT 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 GTCGTACAGATGTCACGAGCTCCTGAATCTCTGTGGAGATCTCAGTATGTTCT 4
QY 541 TCA 543
DB 3 TCA 1
RESULT 2
BY733578 LOCUS
DEFINITION diencephalon Mus musculus cDNA clone G630052E12 5', mRNA sequence.
ACCESSION BY733578
VERSION BY733578
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
AUTHORS Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.

```

```

Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojohori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chochia, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kawai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numa, K., Okido, T., Pavan, W. J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ranaichandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultan, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wyrshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, S. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno,
H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
Nunazaki, R., Ohno, M., Osato, N., Saito, R., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Shiraki, M., Takeda, Y.,
Waki, K., Watanaka, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
Location/Qualifiers
1. .670
/organism="Mus musculus"
/mol_type="mrna"
/strain="G57B/60"
/db_xref="taxon:10090"
/clones="G630052E12"
/sex="male"
/tissue_type="dienecephalon"
FEATURES
source

```

/dev. stages="16 days neonate" /clone.lib="RIKEN full-length enriched, 16 days neonate male diencephalon"		205 a	140 c	169 g	155 t	1 others	
BASE COUNT	ORIGIN						
Query Match	42.9%;	Score	405;	DB	14;	Length	670;
Best Local Similarity	85.6%;	Pred. No.	3.2e-79;				
Matches	450;	Conservative	0;	Mismatches	76;	Indels	0;
Gaps 0;							
QY	1	ATGGAAGTATGAAAATTAGCTAAGACTCGAGAGGGCTTTATGGGGTTGTATTCAAA	60				
DB	145	ATGGAAGTATGAAAAGCTAGCTAAGACTCGAGAGGGCTTTATGGGGTTGTATTCAAG	204				
QY	61	TGCAAGAAACAAACCTCTGGCAAGCTAGTACTGTGTAAAAAATTGTGGAATCTGAAGAT	120				
DB	205	TGCAAGAAACAAATCTTCGGACAGTAGTAGCGATCAAAAATTCGTGGAATCTGAAGAT	264				
QY	121	GATCTGTGTAGAAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAACAT	180				
DB	265	GATCTGTGTGTAGAAAATAGCCCTCGGGAATCCGTATGCTGGAAGCAGTTGAACAC	324				
QY	181	CCAAATCTGTGAACCTCATGAGTGTTCAGGAGAAAGGAAATGCAATTTAGTTT	240				
DB	325	CCAAATCTGTGAACCTCATGAGTGTTCAGGAGAAAGGAAATGCAATTTAGTTT	384				
QY	241	GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGATGTGTGAT	300				
DB	385	GAGTACTGTGATCATACACTGTAAACGAGCTGGAGAGAAACCCAAACGGAGTTCTGAT	444				
QY	301	GGAGTGATCAAAAGCGTATTATGCAACACTTCAAGCTCTTAACTTCTGTCAATACAT	360				
DB	445	GGAGTGATTAAGAGTGTCTATGCGAACCCTTCAAGCCCTTAATCTGTGTCAAGCNC	504				
QY	361	AACTGTATTCAGAGATATAAACCTGAAATATTTCTAATACTAGCAAGGAATTAATC	420				
DB	505	AACTGTATTCAGAGATATAAACCTGAAATATTTCTAATACTAGCAAGGAATTAATC	564				
QY	421	AAATTTGTGACTTCGGTTCACAAATTCGTATTCAGAGAGATGCTACACCGATTAT	480				
DB	565	AAATTTGTGACTTCGGTTCACAAATTCGTATTCAGAGAGATGCTACACCGATTAT	624				
QY	481	GTAGTACGAGATGTACGAGCTCTCAACTCTTGTGGGAGATA	526				
DB	625	GTGACACAGGTGTACGAGCCNCAACTTCTGTGGGAGACA	670				
RESULT 3							
LOCUS	AK016781						
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493341017 product:cyclin-dependent kinase-like 1 (CDC2-related kinase), full insert sequence.						
ACCESSION	AK016781						
VERSION	AK016781.1						
KEYWORDS	HTC; CAP trapper						
SOURCE	Mus musculus (house mouse)						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
AUTHORS	Carninci, P. and Hayashizaki, Y.						
TITLE	High-efficiency full-length cDNA cloning						
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)						
MEDLINE	99279253						
PUBMED	10349636						
REFERENCE	2						
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.						
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes						
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)						
MEDLINE	20499374						

1042:59	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, K., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yonekura, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.					
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer						
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)						
MEDLINE	20530913						
PUBMED	11076861						
REFERENCE	4						
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Cabavant, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anjo, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Flechner, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Mochizuki, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.						
TITLE	Functional annotation of a full-length mouse cDNA collection						
JOURNAL	Nature 409 (6821), 685-690 (2001)						
MEDLINE	21085660						
PUBMED	11217851						
REFERENCE	5						
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.						
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs						
JOURNAL	Nature 420, 563-573 (2002)						
REFERENCE	6 (bases 1 to 1691)						
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyata, T., Yamamura, T., Yaeunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.						
TITLE	Direct Submission						
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)						
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was prepared with a primer [5'-GAGAGAGAAGGATCCAGAGCTCTTTTCTTTTCTTTT-3']. cDNA was						

QY 945 T 945

Db 176 T 176

RESULT 6

BY752739

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY752739 689 bp mRNA linear EST 17-DEC-2002
 cDNA clone F930014106 5', mRNA sequence.
 BY752739
 BY752739.1 GI:27183802
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.
 1 bases 1 to 689;
 Okazaki, Y., Furuno, X., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 N. Kaido, I., Osato, N., Sato, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogari, A., Schenbach, C.,
 Goyobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
 Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,
 L. E., Cousins, S., Dalla, E., Dragani, I. A., Fletcher, C. E., Forrest,
 A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustinch, S., Hirokawa, N., Jackson, J. J.,
 Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M.,
 King, B. L., Kozagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
 P. A., Maglott, D. R., Yaitais, L., Marchionni, J., McKenzie, L., Miki,
 H., Nagashima, T., Nunata, K., Okido, T., Pavan, W. J., Pertea, G.,
 Pesole, G., Petrovsky, N., Pillai, R. K., Pontius, J. J., Reid, J., Ring,
 Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. C., Reid, J., Ring,
 B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
 X., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
 R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
 Macanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sasaki, K., Sasaki, D., Shibata,
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 1246685
 22354683
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
 S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
 Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno,
 H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
 Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,
 Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, Y., Takeda, Y.,
 Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct
 Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Kirk W. Beisel (Boys Town National
 Research Hospital 555 North 30th Street Omaha, NE 68131 USA) whose
 assistance we gratefully acknowledge.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 Location/Qualifiers
 1. 689
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="F930014106"
 /tissue_type="inner ear"
 /dev_stage="adult"
 /clone_lib="RIKEN full-length enriched, adult inner ear"
 213 a 133 c 171 g 168 t 4 others
 BASE COUNT
 CRIGIN
 Query Match 38.1%; Score 360.2; DB 14; Length 689;
 Best Local Similarity 82.7%; Pred. No. 2.7e-69;
 Matches 421; Conservative 0; Mismatches 87; Indels 1; Gaps 1;
 QY 1 ATGGAAGATGATGAAAATTTAGCTAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60
 DB 169 ATGGAAGATGATGAAAAGCTAGCTAGATCGGAGAGGGTCTTATGGGTTGTATTCAG 228
 QY 61 TGCAGAACCAAACTCTGGCAAGTAGTAGCTGTAAAAAATTTTGGGAATCTGAAGAT 120
 DB 229 TGCAGAACCAAACTCTTCGACAAAGTAGTAGCCATCAAAAATTCGTGGAAATCTGAAGAT 288
 QY 121 GATCCTGTGTTAAGAAAATAGCACTAAGAGAAATAGTATGTTGAAGCAATTAACACAT 180
 DB 289 GATCGTGTGTTAGAAAATAGCCCTCGCGGAAATCGTATGCTGAAGCAGTTGAAACAC 348
 QY 180 CCAATCTGTGACCTCATCGAGGTCTCAGAGAAAGAGGAATGCAATTTAGTTT 240
 DB 349 CCAACCTCTGTAACCTCATTCAGGTGTTTCAAGAAAGAGGAGATGCAATTTAGTTT 408
 QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGAGTTGCTGAT 300
 DB 409 GAGTACTGTGATCACACGCTGTTAAACGAGCTGGAGAGAAACCCAAACGAGTTTCTGAT 468
 QY 301 GGAGTGATCAAGAGCTATTATGCAACACTTCAAGCTCTTAATTTCTGCTATATACAT 360
 DB 469 GGAGTGATTAAGAGCTGCTATGCAACACCTTCAAGCCCTTAACTTCTGCTCAAGACAC 528
 QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
 DB 529 AATGTATTATCGGATGTAAACCTGAAACATCTCTAATAACCAAGCAGGGATGANT 588
 QY 420 AAGATTGTGACTTCGGGTTTGCAAAATTTCTGATTCAGAGATGCTTACACCGATTAT 480
 DB 589 AAGATTGTGACTTTGGAATTTGCAAGAAATTTCTAATTNCA-GAGACGCTACACANGACTA 647
 QY 481 GTAGCTAGGAGATGTTACCGAGCTCTGTA 509
 DB 648 TGTCCANCCAGTGTACCGAGCCCCCGA 676
 RESULT 7
 A1385966
 LOCUS
 DEFINITION
 A1385966 500 bp mRNA linear EST 27-JAN-1999
 ml34h09.y1 Stratagene mouse testis (#937308) Mus musculus cDNA
 clone IMAGE:513953 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN

KINASE KIALRE (HUMAN); mRNA sequence.

ACCESSION A1385966
 VERSION A1385966.1 GI:4199429
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
 1 (bases 1 to 500)
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
 B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter,
 E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:307801
 This read is a RESEQUENCE of a previously sequenced mouse clone
 correct orientation
 Seq primer: -40RP from Gibco
 High quality sequence stop: 339.

FEATURES

Location/Qualifiers

1..500

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CD-1"

/db_xref="taxon:10090"

/clone="IMAGE:513953"

/sex="males"

/tissue_type="testis"

/dev_stage="10-12 week old"

/lab_host="SOLR (kanamycin resistant)"

/clone_lib="Stratagene mouse testis (#937308)"

/note="Organ: testis; Vector: pBluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. Average insert size: 2 kb; Uni-ZAP XS Vector;
 -5' adaptor sequence: 5' GAATTCGGCCGAG 3' -3' adaptor
 sequence: 5' CTCGAGTTCCTTTTCTTTT 3'."

BASE COUNT 145 a 114 c 121 g 120 t

ORIGIN

Query Match 35.3%; Score 333.6; DB 9; Length 500;
 Best Local Similarity 82.3%; Pred No. 2e-63;
 Matches 408; Conservative 0; Mismatches 84; Indels 4; Gaps 2;

QY 135 GAAATAGCACTAAGAGAATACGTATGTTGAAGCAATTAACAATCCAAATCTGTGAA 194
 DB 1 2 GAAATAGCCCTCGCGGAATCCGTATGCTGAAG---TTGAACACCAACCAACCTCGTGA 58

QY 195 CCTCATCGAGGTGTTTCAGGAGAAAGGAATGATGTTTGTGATCTGTGATCA 254
 DB 59 CCTCATCGAGGTGTTTCAGGAGAAAGGAATGATGTTTGTGATCTGTGATCA 118

QY 255 TACACTTTTAATGAGCTGGAAGAAACCAATGGAGTTCGTGATGGAGTGCACAAAG 314
 DB 119 CACACTGTTAAAGAGCTGGAGAAACCAACCGAGTTCCTGATGATGATTAAAG 178

QY 315 CGTATTATGCAAAACCTTCAAGCTCTTAAATTTCTGTCATATACATACTGATTTCAG 374
 DB 179 TGTGCTATGCAAAACCTTCAAGCCCTTAACCTTCTGTCACAGCACAATTTGATTTCAG 238

QY 375 AGATATAAAACCTGAAATATTTCTTAATACTAAGCAAGGAATAATCAAGATTGTGACTT 434
 DB 239 GGAATGAAACCTGAAACATCTTAATACCAAGCAAGGATGATTAAGATTGTGACTT 298

QY 435 CGGTTTCACAAAATCTGATTCCAGGAGATGCTACACCGATTATCTAGCTACGAGATG 494
 DB 299 TGGATTTCACGAAATCTTAATTCAGGAGAGCGCTACACAGACTATGTTGCCACGAGTG 358
 QY 495 GTACCGAGCTCTCAAACTTCTTGTGGAGATACATCATGATATGTTTCTTCACTCATATG 554
 DB 359 GTACCGAGCTCTCAAACTTCTTGTGGAGAGACACGAAGTAGCGGTTCCTCTGTAGACGTTG 418
 QY 555 GGCTATTGGTTGTCTTTTTCAGAGACTCTCTGACAGCCAGCCACTGTGGCTCGAATC 614
 DB 419 GGCGTCGCGCTGTGTTATGAGAGACTTCTGACGTGACCCACTCTGGCCCGG-AAATA 477
 QY 615 AGATGTGACCAACTT 630
 DB 478 CGACGTGGACCACTT 493

RESULT 8
 BU221831
 LOCUS
 DEFINITION 824 bp mRNA linear EST 25-NOV-2002
 603750354F1 CSEQCHN04 Gallus gallus cdna clone CHEST661113 5', mRNA
 sequence.
 ACCESSION BU221831
 VERSION BU221831.1 GI:25410266
 KEYWORDS Gallus gallus (chicken)
 SOURCE Gallus gallus
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 824)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 22335534
 12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1OD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers

1..824

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHEST661113"

/tissue_type="whole embryo"

/dev_stage="20-21"

/lab_host="DH10B"

/clone_lib="CSEQCHN04"

/note="Organ: whole embryo; Vector: pBluescript II KS(+);
 Site 1: EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with EcoRI
 , size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaudo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

BASE COUNT 192 a 227 c 239 g 166 t

ORIGIN

The WashU-HHMI Mouse EST Project	
Unpublished	
Contact: Marra M/Mouse EST Project	
WashU-HHMI Mouse EST Project	
Washington University School of MedicineP	
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
Tel: 314 286 1800	
Fax: 314 286 1810	
Email: mouseest@wustl.edu	
This clone is available royalty-free through LLNL ; contact the	
IMAGE Consortium (info@image.llnl.gov) for further information.	
XGI:307801	
Seq primer: -28ml3 rev1 ET from Amersham	
High quality sequence stop: 442.	
Location/Qualifiers	
1..938	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="CD-1"	
/db_xref="taxon:10090"	
/clones="IMAGE:513953"	
/sex="males"	
/tissue_type="testis"	
/dev_stage="10-12 week old"	
/lab_host="SOLR (kanamycin resistant)"	
/clone_lib="Stratagene mouse testis (#937308)"	
/note="Organ: testis; Vector: pBluescript SK-; Site 1:	
ECOR1; Site 2: Xho1; Cloned unidirectionally. Primer:	
Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;	
-5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor	
sequence: 5' CTCAGTCTTTTCTTTTCTTTT 3"	
BASE COUNT	254 a 215 c 236 g 233 t
ORIGIN	
Query Match	33.5%; Score 316.8; DB 9; Length 938;
Best Local Similarity	82.9%; Pred. No. 1.2e-59;
Matches	398; Conservative 0; Mismatches 77; Indels 5; Gaps 3;
QY	135 GAAATATACACTTAAGAGAAATACGTATGTTGAAGCAATTAACATCCAAATCTGTGAA 194
DB	1 GAAATATACCTCGGGAAATCCGTATGCTGAG--TTGAAACACCAACCTCGTGA 57
QY	195 CCTCATCGAGGTGTCAGGAGAAAGAAAGCAATTCATTTAGTTTTCGAATCTGTGATCA 254
DB	58 CCTCATCGAGGTGTCAGAGAAAGAAAGCAATTCATTTAGTTTTCGAATCTGTGATCA 117
QY	255 TACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGATGGAGTATCAAAAG 314
DB	118 CACACTGTTAAACGAGCTGGAGAGAAACCCCAACGGAGTTTCTGATGGAGTATTAAG 177
QY	315 CGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACATAACTGTATTACAG 374
DB	178 TGTGCTATGGCAACCCCTTCAAGCCCTTAACTTCTGTCAAGCAACAAATGTATTCTATCG 237
QY	375 AGATATAAACTCGAAATATTTCTAATACTAAGCAAGGAATTAATCAAGATTTGTGACTT 434
DB	236 GGATGTAAACCTGAAACATCTCTAATAACCAAGCAAGGATGATAAGATTTGTGACTT 297
QY	435 CGGTTTGCACAAATTCGTGATTCAGGAGATGCTTACCCGATTTATGCTAGCTACGAGATG 494
DB	298 TGGATTTGCACGAATTTCTAATTTCCAGGAGACGGCTACACAGACTATGTTGCCACCGAGTG 357
QY	495 GTACCGAGCTCTCAACTCTTGTGGGAGATCTCAGTATGGTTCTTTCAGTCCGATATATG 554
DB	358 GTACCGAGCCCCCAACTTCTCGTGGGAGACAGAGTACGGTTCTCTGTAGACGTGTG 417
QY	555 GCGTATTGGTTGTGTTTTCAGAGACTCTTCAGAGCGCCAGCCACTGTGGCTCGGAAATC 614
DB	417 GCGCTCGGCTGTGTTTTCGAGA-CTCTCTGACGGGTGAG-CACTCTGCGCCGGGAAAC 475
RESULT 10	
AV986182	

Query Match	33.8%; Score 319.4; DB 13; Length 824;
Best Local Similarity	67.2%; Pred. No. 3e-60;
Matches	483; Conservative 0; Mismatches 231; Indels 5; Gaps 2;
QY	40 TCTATTGGGGTGTATTCAAATGCGAGAAACAAACCTCTGGACAGTAGTAGCTGTATAA 99
DB	2 TCTTACGGCGTGTCTTCAAGTCGGCAACAGGAGACGGGGAGATCGTGGCCATCAAG 61
QY	100 AAATTTGTGGAAATCGAAGATGATCTGTGTTTAAGAAATAGCACTAAAGAGAAATACGT 159
DB	62 AAGTTCTCGAGTCCGAGAGGACCCGGTGATCGGAAGATCGCGCTCGCGGAGTCCGC 122
QY	160 ATGTTGAAGCAATTAACATCCAAATCTTGTGAACCTCTACAGAGTGTTCAGAGAAA 219
DB	122 ATGCTGAAGCAACTGAACACCCCAACCTGTGTAACCTGTGGAGTGTTCAGAGAAAG 181
QY	220 AGGAAATGCAATTTAGTTTGAATACCTGTGATCATACATTTAAATGAGCTGGAAAGA 279
DB	182 AGGAAGCTGACCTGGTCTTTGAGTACTGTGACCAACCGTTCCTCAGGAGCTGGACAAG 241
QY	280 AACCCAAATGGAGTGTCTGTATGGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCT 339
DB	242 CACCCCGGGGGTCCCGGACGACGTGTGAGGAGCATTACCTGGCAGACCTCCAACT 301
QY	340 CTTAATTTCTGTATATACATACTGATTTACACAGAGATATAACCTGAAATATCTTA 399
DB	302 GTGAACCTTTTGTCAACACAACTGCTATCCATCGAGATGTAAAGCCAGAAACATCTCG 361
QY	400 ATAACTAAGCAAGCAATTAATCAAGATTTGTGACTTCGGGTTCGACAAATTTCTGATT 456
DB	362 ATAAAGCAAGCACTCCATCATCAAACTCTGTGACTTCGGATTTCTGCGATCTGACTGGC 421
QY	457 CCAGGAGATGCTACACGATATGTAGTACAGAGTGGTACCGAGCTCCTGAACCTTCT 516
DB	422 CCAGGTGATTACTACACAGACTAGTGGCAACAGGTGTTACCGCTCCCGACAGCTGCTG 481
QY	517 GTGGGAGATCTCAGTATGGTTCTTCAGTCGATATATGGCTATTGTTGTTTTCGA 576
DB	482 GTGGGGGACACGAGTATGGCCCTCTGTGGACGTGTGGCAATAGGCTGTGCTTTGGC 541
QY	577 GAGCTCTCAGAGCCAGCACCTGTGGCTTGGAAATCAGATGTGGACCAACTTTATCTG 636
DB	542 GAGCTGCTCTCGGGGTGCCACTGTGGCCGGCAAGTGCAGCGTGGACAGCTGTACCTC 601
QY	637 ATAATCAGAACTAGGAAATTAATCCCAAGACATCAATCAATCTTTAA--AAGTAACG 694
DB	602 ATCCGAGAACATTTGGGGATCTTATTCCAGGCCAGCAAGTGGTTTCAGCACCAAAAC 661
QY	695 GGTTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTTTGAGGAAAG 753
DB	662 AGTTCTTCAGCGGTAAACGATTCAGACCCAGAGAGCATGGAGCCATTTGGGCAATG 722
RESULT 9	AA061797
LOCUS	ml34h09.rl Stratagene mouse testis (#937308); Mus musculus cDNA
DEFINITION	clone IMAGE513953 5', similar to gb:X66358 SERINE/THREONINE-PROTEIN
ACCESSION	KINASE KTIALRE (HUMAN); mRNA sequence.
VERSION	AA061797
KEYWORDS	AA061797.1 GI:15555606
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
	1 (bases 1 to 938)
	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
	Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
	Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
	Theisinger, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
	Waterston, R.

LOCUS AV986182 598 bp mRNA linear EST 14-MAR-2002
 DEFINITION AV986182 Nori Satoh unpublished cDNA library, larva Ciona
 accession AV986182 intestinalis cDNA clone cili4b18 5', mRNA sequence.
 VERSION AV986182
 KEYWORDS EST.
 SOURCE BP016186.1 GI:19475050
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.
 REFERENCE 1 (bases 1 to 598)
 AUTHORS Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
 TITLE Expressed genes in Ciona intestinalis
 JOURNAL Unpublished
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satohgascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 FEATURES
 Source
 1..598
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cili4b18"
 /tissue_type="whole animal"
 /dev_stage="larva"
 /clone_lib="Nori Satoh unpublished cDNA library, larva"
 BASE COUNT 179 a 105 c 149 g 165 t
 ORIGIN
 Query Match 32.5%; Score 306.8; DB 9; Length 598;
 Best Local Similarity 70.7%; Pred. No. 1.8e-57;
 Matches 423; Conservative 0; Mismatches 172; Indels 3; Gaps 1;
 QY 11 ATGMAAATTACTGACTGAGGAGGCTCTATCGGGTGTATTCCAAATCGAACA 70
 DB 1 ATGAGAAGATAGGTAAAGATAGGGGAGGATCTATGGCATTTGTATTAATGTCGGAACC 60
 QY 71 AAACCTCTGGCAAGTAGTAGCTGTATAAAATTTGTGGAAATCTGAAGATGATCCTGTG 130
 DB 61 GTGACTCGGACAAATTGTGGCCATCAAAAATTTGTGAGTCAGAGATGACCCGTGA 123
 QY 131 TTAAGAANTACACTAGAGAAATACGTATGTTGAAGCAATTAACATCCAAATCTTG 190
 DB 121 TCAAGAAATTCGCTGAGGGAATCAGAATGTTAAAGCAATTAAGACCAATTAACCTGG 180
 QY 191 TGAACCTCATGAGGTGTTTCAGGAGAAAAGGAAAATGCATTTAGTTTGAATACCTGTG 250
 DB 181 TGAACTGATGAAAGTTTCCGCGAAACGCAATTAACATCTGTGTTGAGGTATGTG 240
 QY 251 ATCATACATTTTAAATAGCTGGGAAGAACCCCAATGGAGTGTGATGGAGTATCA 310
 DB 241 ATCATACATTTTAAATAGTTTGGCAAAACACATCAGAGGAGTACCAGAACATCATGTA 300
 QY 311 AAAGGTATTTATGGAAACACATCTAAGCTCTTAATTTCTGTCATATACATAACTGATTC 370
 DB 301 AAGTATTTGGAAGTTTCTTAAGCTGTTTCACTTTTGGCATCAGCATTAATTTGATAC 360
 QY 371 ACAGAGATATAAACCTGAAAATTTCTAATACTAAGCAAGGAATTAATCAAGATTTGTG 430
 DB 361 ATCGTATGTAAGAACCCAGAGACATATTAACTCACTAAGCAAGGTGTTATCAAGTTGTG 420
 QY 431 ACTTCGGGTTTGCACAAATCTGA--TTCCAGGAGATGCCCTACACCGATTAATGATCA 487
 DB 421 ACTTTGGATTTGCAAGAAATTTTAACTGTGCTCTGGGACGATTACACCGACTATGGCCA 480
 QY 488 CGAGATGGTACCGAGCTCTGAACTCTTTGTGGGAGATACCTCAGTATGGTTCTTCAGTCG 547
 DB 481 CTAGGTGGTACAGAGCACCTGAGCTCTCTGTCGGGACACTCAGTATGGCCCTCCAGTGG 540

QY 548 ATATATGGCTATTGGTTGTGTTTTTCAGAGCTCTTGACAGGCCAGGCACCTGTGGCC 605
 DB 541 ATGTGTGGCAATAGGTTCGCTATTTCGAGTTGTTGCGGACAGGCAATTATGGCC 598
 RESULT 11
 LOCUS BP016186 597 bp mRNA linear EST 15-MAR-2002
 DEFINITION BP016186 Nori Satoh unpublished cDNA library, young adult Ciona
 intestinalis cDNA clone ciad64e08 5', mRNA sequence.
 VERSION BP016186
 KEYWORDS EST.
 SOURCE BP016186.1 GI:19507663
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.
 REFERENCE 1 (bases 1 to 597)
 AUTHORS Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
 TITLE Expressed genes in Ciona intestinalis
 JOURNAL Unpublished
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satohgascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 FEATURES
 Source
 1..597
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="ciad64e08"
 /tissue_type="whole animal"
 /dev_stage="young adult"
 /clone_lib="Nori Satoh unpublished cDNA library, young
 adult"
 BASE COUNT 184 a 100 c 139 g 173 t 1 others
 ORIGIN
 Query Match 30.6%; Score 289.2; DB 12; Length 597;
 Best Local Similarity 71.0%; Pred. No. 1.4e-53;
 Matches 397; Conservative 0; Mismatches 159; Indels 3; Gaps 1;
 QY 2 TGGAAAGTTTGAATAATAGCTAGAGTGGAGAGGCTCTTATGGGTTGTATTCAAT 61
 DB 38 TGGAGAAATATGAGAAGATAGGTAAGATAGGGGAGGATCTTATGGCATTTGTGTTAAAT 97
 QY 62 GCAGAAACAAAACCTCTGCAAGTAGTAGCTGTATAAAATTTGTGGAATCTGAAGATG 121
 DB 98 GTCGGAACCGTGCATCGGACAAATTTGCGCCATCAAAAATTTGTCGATCAGAAATG 157
 QY 122 ATCTCTGTTTGAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACATC 181
 DB 158 ACCCCCTGATCAAGAAAATCGCGTGGAGGAAATCAGAAATGTTAAAGCAATTTGAACACA 217
 QY 182 CAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAATCGATTTAGTTTGTG 241
 DB 218 ATAACTTGGTGAACCTGATTAAGTTTTCGTCGAAAACGCAAAATTAACATCTTGTGTG 277
 QY 242 AATACTGTGATCATACTCTTTTAAATGAGCTGGAAAGAAAACCCCAATGGAGTTGCTGATG 301
 DB 278 AGTATTGTGATCATACAGTTTAAATGAATTTGGAACAAACATCAGAGGAGTACCAGAAC 337
 QY 302 GAGTGATCAAAAGCGTATTATGGCAACACATTCAGCTCTTAATTTCTGTCATATACATA 361
 DB 338 ATCATGTTTAAACGATTTCTTTGGCAAGTTCTTCAAGGTGTTCACTTTTGGCCATCAGCATA 397
 QY 362 ACTGTATTTCAGAGATATAAAGCTGAATATTTCTTAATACTTAAGCAGGAAATATCA 421
 DB 398 ATTGTATACATCGTGATGTAAACAGAGAACATATTAACTCACTAAGCAAGGTGTTATCA 457

```

QY 422 AGATTGTGACCTCGGGTTTSCACAAATCTTGA---TTCAGGAGATGCTTACACCGATT 478
Dh 458 AGTTGTGTGACTTGTGATTTCGAGAATTTTAACTGTCTCGGCGGAGATTACACCGACT 517
QY 479 ATGTAGCTACAGATGTTACCGAGCTCTGAACTCTTCTGTGGGAGATCACTCAGTATGTT 538
Dd 518 ATGTGGCCACTAGGTGTACAGAGACCTGAGCTCTCTGTGGGAGACACTCAGTATGGCC 577
QY 539 CTTCAGTCGATATGGGC 557
Dd 578 CTNAGTGGATGTGTGGGC 596

RESULT 12
LOCUS CB400506 565 bp mRNA linear EST 15-MAY-2003
DEFINITION OSF177F:0.1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
ACCESSION CB400506
VERSION CB400506.1 GI:30742233
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Phaditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 565)
Rebout,J., Vaglio,P., Rual,J.F., Jaresch,P., Martinez,M., Armstrong
,C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T., Hudson
,J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
Endress,G.A., Jenna,S., Chevet,E., Papasotiropoulos,V., Tolias,P.P.,
Ratsek,J., Snyder,R., Huang,R., Chance,M.R., Jee,H.,
Doucette-Stamm,L., Hill,D.E. and Vidal,M.
C. elegans ORFeome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
Nat. Genet., (2003) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFeome
cloning project : Contact david_vidal@dfci.harvard.edu or
marc_vidal@dfci.harvard.edu
POLYA=No.

FEATURES
source Location/Qualifiers
1..565
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmcDNA"
/note="The AD-wrmcDNA library was generated with poly(A)-
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pPCR86"
BASE COUNT 184 a 88 c 135 g 158 t
ORIGIN

Query Match 28.6%; Score 269.9; DB 14; Length 565;
Best Local Similarity 67.6%; Pred. No. 2.8e-49;
Matches 379; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 31 CGAGAGGGTCTTATGGGTTGTATTCCAAATCAGAACAAACCTCTGACCAAGTAGTA 90
Dd 5 GGTGAAGGTTTCATATGGAGTTGTATATAAATGTAAAAATAGGGATACCTGGACAAATGTG 64

```

```

QY 91 GCTGTTAAAAAATTTGTGGAATCTGAAGATGATCCTGTGTTTAAGAAAATAGCACTAGA 150
Dd 65 GCAATCAAAAAGTTTGTGAAACTGAGATGATCGCATATTAAGAAAATTTGCCCTACGG 124
QY 151 GAAATACGTATGTTGAACCAATTAACACATCCAAATCTTGTGAACCTCATCCAGGTGTT 210
Dd 125 GAAATCAGATGCTGAAGCAACTGAACATCAAAATTTGGTTGGATTGATTGAAGTGTTC 184
QY 211 AGGAGAAAAGGAAATGCAATTTAGTTTGAATCTGTGATCATACACATTTTAAATGAG 270
Dd 185 AAACGGAACCGGAAACITTCATCTGCTTTTGAATCTTGTGATCGAATGTACTTCCAGAG 244
QY 271 CTGGAAAGAAACCCAAATGGAGTTGCTGATGGAGTATCAAAAGCGTATTATGCGCAACA 330
Dd 245 TTGGAGAGAATCCGATGGAGTTAAGCATCACTCATAAAGAAAATAATTTATCAATTG 304
QY 331 CTTCAAGCTCTTAATTTCTGTCATATACATAACTGTATTACAGAGATATAAAACCTGAA 390
Dd 335 STAGAGCTCTTAAATCTGTACAGTCAAAATGTATTTCATCGAGATGTGAACCCGAA 364
QY 331 AATATTCTAATAACTAAGCAAGGAATATCAAGATTTGTGACTTCGGGTTTGCAACAATT 450
Dd 365 AACATTTCTTGACACGGAATGATCAAGTGAACTTGAGATTTCGGATTTCGGAATA 424
QY 451 CTGATCCAGGAGATGCTACACCGATTATGTAGCTACGAGATGTTACCGAGCTCTCTGAA 510
Dd 425 ATAAACACGACGGAATGTATCTGACTATGTGCAACTCTGTGTTACCGTAGTCCAGAA 484
QY 511 CTTCTTTGGGANGATCACTACATATGTTCTTCAGTCAATATATGCGGTATGGTGTGTT 570
Dd 485 CTTCTGGTGGTGTGATGTTCAATATGGGCCACCTGTAGACATTTGGGCTGTAGGATGTGA 544
QY 571 TTTCGAGAGCTCTCTGACAGGC 591
Dd 545 TATCGGAGCTGTTGACTGGC 565

RESULT 13
LOCUS AW106692 731 bp mRNA linear EST 20-OCT-1999
DEFINITION IMAGE-2216229.5, similar to gb:X66358 SERINE/THREONINE-PROTEIN
KINASE KTA13E (HUMAN);, mRNA sequence.
ACCESSION AW106692
VERSION AW106692.1 GI:6077492
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 731)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:1006441
Seq primer: custom primer used
High quality sequence stop: 433.
Location/Qualifiers
1..731
/organism="Mus musculus"
/mol_type="mRNA"

```

FEATURES
source

[illegible]

RESULT 14	CA334835/c	CA334835	645 bp	mRNA	linear	EST C4-NOV-2002
LOCUS						
DEFINITION						NISC l101b12.y1 COGENE 8.5 EPT Homo sapiens cDNA clone IMAGE:5605870 5', mRNA sequence.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 16:54:37 : Search time 73.9079 Seconds
(without alignments)
6216.928 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaagaatgatgaaaatt.....aggatcttcggtcaaaagt 1041

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	251.8	24.2	1701	4	US-09-411-628-5
2	251.8	24.2	3080	4	US-09-411-628-3
3	163.8	15.7	903	2	US-08-874-347-9
4	163.8	15.7	903	3	US-09-093-522-9
5	162	15.6	1476	2	US-08-969-106-1
6	162	15.6	1476	4	US-09-338-125-1
7	162	15.6	1476	4	US-09-266-225D-13
8	162	15.6	1635	4	US-09-417-197-112
9	162	15.6	1635	4	US-09-417-197-114
10	154.2	14.8	1050	4	US-09-220-132-3
11	134.8	12.9	1825	4	US-09-620-312D-313
12	133.2	12.8	1089	1	US-08-154-915-1
13	133.2	12.8	1089	2	US-08-464-517-37
14	133.2	12.8	1089	2	US-08-246-361A-37
15	133.2	12.8	1089	3	US-08-463-772-37
16	133.2	12.8	1089	5	PCT-US93-09945-1
17	128	12.3	1002	1	US-08-463-090B-3
18	116	11.1	1070	1	US-08-463-090B-5
19	112	10.8	1308	4	US-09-801-861-4
20	112	10.8	2203	4	US-09-801-861-1
21	103	9.9	2747	2	US-08-874-347-1
22	103	9.9	2747	3	US-09-093-522-1
23	100.6	9.7	2467	4	US-09-206-344A-3
24	96.2	9.2	1673	3	US-09-347-801-15
25	94.2	9.0	1272	4	US-09-206-344A-1
26	94.2	9.0	1549	4	US-09-206-344A-13
27	89.6	8.6	90541	4	US-09-759-359A-3

C	28	87.4	8.4	6340	1	US-08-187-785-3	Sequence 3, Appli
	29	87.4	8.4	87350	3	US-08-781-891-79	Sequence 79, Appl
	30	87.4	8.4	87350	4	US-09-618-166-79	Sequence 79, Appl
	31	87.4	8.4	87543	4	US-09-793-211-3	Sequence 3, Appli
	32	87.4	8.4	99500	4	US-09-798-096-10	Sequence 10, Appl
	33	87.2	8.4	1946	4	US-09-620-312D-462	Sequence 462, App
	34	87.2	8.4	70000	4	US-09-851-896-3	Sequence 3, Appli
C	35	87	8.4	8779	2	US-08-750-703-4	Sequence 4, Appli
	36	86.6	8.3	2624	1	US-08-032-382B-1	Sequence 1, Appli
	37	86.4	8.3	6804	4	US-09-740-041-3	Sequence 3, Appli
	38	86	8.3	72604	4	US-09-268-992-7	Sequence 7, Appli
	39	86	8.3	72604	4	US-09-657-474-7	Sequence 7, Appli
	40	85.6	8.2	604	4	US-09-659-845A-22	Sequence 22, Appl
C	41	85.6	8.2	16595	4	US-09-146-053-7	Sequence 7, Appli
C	42	85.4	8.2	7216	2	US-08-750-703-3	Sequence 3, Appli
	43	85.4	8.2	112132	4	US-09-741-150-3	Sequence 3, Appli
C	44	85.2	8.2	18853	4	US-09-820-005-3	Sequence 3, Appli
	45	84.6	8.1	1249	4	US-09-220-132-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-411-628-5
; Sequence 5, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; EARLIER FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-09-411-628-5

Query Match	24.2%	Score	251.8;	DB	4;	Length	1701;
Best Local Similarity	62.9%	Pred. No.	9.1e-61;				
Matches	407;	Conservative	0;	Mismatches	237;	Indels	3;
Gaps	1;						
Qy	1	ATGGAAAGTATGAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGATTCAA	60				
Db	1	ATGGAAATATGAGACCTTGGATTGGTGGAGAGGGAGTTATGGATGGTATGAG	60				
Qy	61	TCGAAACAAACCTCTGGACAAGTAGTACGTGTAAAAAATTTTGGAAATCTGAAGAT	120				
Db	61	CTAGGAATAAAGATAGTGGAGAAATTCGCCATCAAGAGTTCTCTAGAAAGTGATGAT	120				
Qy	121	GATCTGTGTAGAAATAGCCTAAGACAAATACGTATGTTGAAGCAATTTAAACAT	180				
Db	121	GACAAATGGTTTAAAAAAATTTGCTATCGGAGAAATCAAGTTTACTAAAGCACTAGGCAT	180				
Qy	181	CCAAATCTTGTGAACCTCATCGAGTGTTCAGGAGAAAAAGAAAAATGCAATTTAGTTTTT	240				
Db	181	GAATATTTGGTGAATCTGTGGAGGTGTAAAAAANAACGATGGTACCTAGTCTTT	240				
Qy	241	GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT	300				
Db	241	GAATTTGTGGACCAACCACTTCTTGAGCTTGGAACTCTTTCCAAATGGACTAGATGAC	300				
Qy	301	GGAGTGTCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATACAT	360				
Db	301	CAAGTAGTCAAAAGTATTCTTTTTCAGATTATTAATGGAATTTGATTTTGCACAGTCAC	360				
Qy	361	AACTGTATTCACAGATATAAACCCTGAAATATTTCTAATAACTAAGCAAGGAATATC	420				

Db 361 AATATCATACATAGATATAAAGCCAGAGATATATTTGGTCTCCAGTCTGGCGTTGTC 420
QY 421 AAGATTCTGTGACTTCGGGTTTGGACAAATTCG---ATTCCAGAGATGCTTACACCGAT 477
Db 421 AAGTTATGTGATTTGGATTTGACAGGACACTGGCAGCTCCCGAGAGGTTTACACTGAT 480
QY 478 TATGTAGCTACGAGATGTCACCGAGCTCCTGAACCTTTCTGTGGGAGATACACTCAGTATGGT 537
Db 481 TATGTGGCAACTCGATGTCACAGAGCTCCAGAACTACTGCTGTGGTATGCAAGTATGGC 540
QY 538 TCTTCAGTTCGATATATGGCTATTTGGTTGTTTTCGAGAGCTCTCCAGAGCCAGCCA 597
Db 541 AAAGCTGTGGATGTGTGGGCCATTTGGTTGCTGGTAACTGAAATGCTCATGGGGGAACCC 600
QY 598 CTGTGGCTCGAAATCAGATGTGGACCAACTTTTATCTGATAATCAG 644
Db 601 CTGTTCTCGAGACTCTGATTTGATGATCAGCTTTTATCTTATTATGAG 647

RESULT 2
US-09-411-628-3
; Sequence 3, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CONA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411-628
; EARLIER FILING DATE: 1999-10-01
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3080
; TYPE: DNA
; ORGANISM: *Orcytolagus cuniculus*
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (249)...(1949)
US-09-411-628-3

Query Match 24.2%; Score 251.8; DB 4; Length 3080;
Best Local Similarity 62.9%; Pred. No. 1.2e-60;
Matches 407; Conservative 0; Mismatches 237; Indels 3; Gaps 1;
QY 1 ATGGAAGATGATGAAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
Db 249 ATGGAAGATATGGAACCTTGGATTGGTGGAGAGGGAGTTATGMAATGGTATCAAG 306
QY 61 TGCAGAAACAAACCTCTGGACAGTAGTACTGTGTTAAATAATTTGGAACTCAGAGAT 120
Db 309 TGTAGGAATAAGATATGAGGAGAAATTTGGCCATCAAGAAATTTCTAGAAAGTATGAT 368
QY 121 GATCTGTGTTTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
Db 369 GACAAAATGTTTAAATAATTTGCTATGGAGAAATCAAGTTACTAAGCAACTAGCGCAT 428
QY 181 CCAATCTTTGGAACCTCATCGAGGTGTCAGGAGAAAGAAATGCAATTTAGTTT 240
Db 429 GAAATTTGGTGAATCTGTTGGAGGTGTGTAATAAATAAATAAAGCACTAGTCTT 488
QY 241 GAATCTGTGATCATACACTTTTAAATGAGTGGAGAAACCCAAATGGAGTTCGTAT 300
Db 489 GAATTTGTGACCAACAGCAATTTCTGATGACTTGGAACTCTTCCAAATGGAGTATGAC 548
QY 301 GGAGTGATCAAAAGCGTATTTATGCAACACCTTCAGCTCTTAATTTCTGTCATATACAT 360
Db 549 CAAGTAGTTCAAAAGTATTTGTTTCAGATTATTAATGAAATGGATTTTGTACAGTAC 608
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATAACTTAAGCAAGSATAATC 420

Db 603 AATATCATACATAGATATAAAGCCAGAGATATATTTGGTCTCCAGTCTGGCGTTGTC 668
QY 421 AAGATTCTGTGACTTCGGGTTTGGACAAATTCG---ATTCCAGAGATGCTTACACCGAT 477
Db 669 AAGTTATGTGATTTGGATTTGACAGGACACTGGCAGCTCCCGAGAGGTTTACACTGAT 728
QY 478 TATGTAGCTACGAGATGTCACCGAGCTCCTGAACCTTTCTGTGGGAGATACACTCAGTATGGT 537
Db 729 TATGTGGCAACTCGATGTCACAGAGCTCCAGAACTACTGCTGTGGTATGTCAGTATGGC 788
QY 538 TCTTCAGTTCGATATATGGCTATTTGGTTGTTTTCGAGAGCTCTCCAGAGCCAGCCA 597
Db 789 AAAGCTGTGGATGTGTGGGCCATTTGGTTGCTGGTAACTGAAATGCTCATGGGGGAACCC 848
QY 598 CTGTGGCTCGAAATCAGATGTGGACCAACTTTTATCTGATAATCAG 644
Db 849 CTGTTCTCGAGACTCTGATTTGATGATCAGCTTTTATCTTATTATGAG 895

RESULT 3
US-08-874-347-9
; Sequence 9, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...900
; OTHER INFORMATION:
US-08-874-347-9

Query Match 15.7%; Score 163.8; DB 2; Length 903;
Best Local Similarity 55.2%; Pred. No. 2.7e-36;
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;

```
QY 1 ATGAAAAGTATGAAAATAGCTAAGAGCTGAGAGGGTCTATGGGTTGTTATCAA 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGAGCAATATCAGAGTTAGAGAGATGAGAGAGAACTATGGAGTGTGTTATAAA 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TGAGAAAACAAACCTCTGAGCAAGTAGTAGTGTGTTAAAAAATTTGGAATCTGAAGAT 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GCAAGGATCTTGAAGAGTGTGCAATTTAGCTCTTAAGAAAATCCGATTAGAGCAGAA 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GATCCTGTTCTTAAGAAATAGCACTAAGAGAAATAGCTATGTCAGACMATTAAGACAT 160
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GATCAGGAGTCTCTAGTACAGCAATCTGAGATATCTACTTTGAGAGAGATGACCAAT 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 CCAATCTTCTGAACCTCATCAGGTGTTTCAAGAGAAAAGGAAAATGATCTTATGTTTT 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GATAATCTTGAAGCTTTTGAATATAATTCATCAAGAGTCAAGTTTATATCTTGTITTT 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 GAATAGTGTATCATACATCTTTAAATGAGCTGGAAGAAACCAATG-----GAGTT 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GAATTTCTTGATCTTGAATTTAAAGATATATGATAGTATTTCCAAAGGACATGTCCT 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 GCTGATGAGTGTATCAAAAGCGTATTATGCGAAACACTTCAAGCTCTTAATTTCTGTCAT 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GGTGCAAGAAATGATTAAGAGTTTATGTCACAACTTGTATCAGGTGTAAATATTGTCAT 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 ATACATACTGTATTCACAGAGATATAAACTGAAATATTTCTAAATCACTAAGCAGGA 414
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TCTCATCGTATTTCTTCATCGTACTTGAAACCAACAAATCTTTATAGATCGAGAGGA 420
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 ATAATCAAGATTTGTGACTTCGGGTTTGCACAAA---TTCTGATTTCCAGGAGATGCTTAC 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 AATCTTAAATACAGATTTTGGGCTTGCAGAGGCGTTTGGTTCCTATGCGTGGTTAT 480
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 472 ACCGATATGAGTACAGAGATGTTACGAGCTCTGAACTCTTTGCGGAGATCTCAG 531
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 AAGCCATTATTTCCAGGCTGATCTGAAATGATGAATATTTAGAAATATTTAGAAATATTA 560
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 G 652
Db 661 G 661
```

RESULT 4
US-09-671-050-9
Sequence 9, Application: US/09091522
Patent No. 6015700
GENERAL INFORMATION:

APPLICANT: Limper, Andrew H.
APPLICANT: Leof, Edward B.
APPLICANT: Thomas, Charles F.
APPLICANT: Gustafson, Michael P.
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
TITLE OF INVENTION: CARINI
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P.C., P.A.
STREET: 60 South Sixth Street, Suite 3300
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

```
QY 1 ATGAAAAGTATGAAAATAGCTAAGAGCTGAGAGGGTCTATGGGTTGTTATCAA 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGAGCAATATCAGAGTTAGAGAGATGAGAGAGAACTATGGAGTGTGTTATAAA 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TGAGAAAACAAACCTCTGAGCAAGTAGTAGTGTGTTAAAAAATTTGGAATCTGAAGAT 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GCAAGGATCTTGAAGAGTGTGCAATTTAGCTCTTAAGAAAATCCGATTAGAGCAGAA 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GATCCTGTTCTTAAGAAATAGCACTAAGAGAAATAGCTATGTCAGACMATTAAGACAT 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GATGAGGAGTCTTGAAGAGTGTGCAATTTGAGCTCTTAAGAAAATCCGATTAGAGCAGAA 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 CCAATCTTCTGAACCTCATCAGGTGTTTCAAGAGAAAAGGAAAATGATCTTATGTTTT 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GATAATCTTGAAGCTTTTGAATATAATTCATCAAGAGTCAAGTTTATATCTTGTITTT 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 GAATAGTGTATCATACATCTTTAAATGAGCTGGAAGAAACCAATG-----GAGTT 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GAATTTCTTGATCTTGAATTTAAAGATATATGATAGTATTTCCAAAGGACATGTCCT 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 GCTGATGAGTGTATCAAAAGCGTATTATGCGAAACACTTCAAGCTCTTAATTTCTGTCAT 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GGTGCAAGAAATGATTAAGAGTTTATGTCACAACTTGTATCAGGTGTAAATATTGTCAT 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 ATACATACTGTATTCACAGAGATATAAACTGAAATATTTCTAAATCACTAAGCAGGA 414
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TCTCATCGTATTTCTTCATCGTACTTGAAACCAACAAATCTTTATAGATCGAGAGGA 420
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 ATAATCAAGATTTGTGACTTCGGGTTTGCACAAA---TTCTGATTTCCAGGAGATGCTTAC 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 AATCTTAAATACAGATTTTGGGCTTGCAGAGGCGTTTGGTTCCTATGCGTGGTTAT 480
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 472 ACCGATATGAGTACAGAGATGTTACGAGCTCTGAACTCTTTGCGGAGATCTCAG 531
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 AAGCCATTATTTCCAGGCTGATCTGAAATGATGAATATTTAGAAATATTTAGAAATATTA 540
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 532 G 532
Db 541 G 541
```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/093,522
FILING DATE: 08-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/874,347
FILING DATE: 13-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055002
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
NAME/KEY: Coding Sequence
LOCATION: 1...900
OTHER INFORMATION:
US-09-093-522-9

Query Match 15.7%; Score 163.8; DB 3; Length 903;
Best Local Similarity 55.2%; Pred. No. 2.7e-36;
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;
QY 1 ATGAAAAGTATGAAAATAGCTAAGAGCTGAGAGGGTCTATGGGTTGTTATCAA 60
Db 1 ATGAGCAATATCAGAGTTAGAGAGATGAGAGAGAACTATGGAGTGTGTTATAAA 60
QY 61 TGAGAAAACAAACCTCTGAGCAAGTAGTAGTGTGTTAAAAAATTTGGAATCTGAAGAT 120
Db 61 GCAAGGATCTTGAAGAGTGTGCAATTTAGCTCTTAAGAAAATCCGATTAGAGCAGAA 120
QY 121 GATCCTGTTCTTAAGAAATAGCACTAAGAGAAATAGCTATGTCAGACMATTAAGACAT 180
Db 121 GATGAGGAGTCTTGAAGAGTGTGCAATTTGAGCTCTTAAGAAAATCCGATTAGAGCAGAA 180
QY 181 CCAATCTTCTGAACCTCATCAGGTGTTTCAAGAGAAAAGGAAAATGATCTTATGTTTT 240
Db 181 GATAATCTTGAAGAGTGTGCAATTTGAGCTCTTAAGAAAATCCGATTAGAGCAGAA 240
QY 241 GAATAGTGTATCATACATCTTTAAATGAGCTGGAAGAAACCAATG-----GAGTT 294
Db 241 GAATTTCTTGATCTTGAATTTAAAGATATATGATAGTATTTCCAAAGGACATGTCCT 300
QY 295 GCTGATGAGTGTATCAAAAGCGTATTATGCGAAACACTTCAAGCTCTTTAATTTCTGTCAT 354
Db 301 GGTGCAAGAAATGATTAAGAGTTTATGTCACAACTTGTATCAGGTGTAAATATTGTCAT 360
QY 355 ATACATACTGTATTCACAGAGATATAAACTGAAATATTTCTAAATCACTAAGCAGGA 414
Db 361 TCTCATCGTATTTCTTCATCGTACTTGAAACCAACAAATCTTTATAGATCGAGAGGA 420
QY 415 ATAATCAAGATTTGTGACTTCGGGTTTGCACAAA---TTCTGATTTCCAGGAGATGCTTAC 471
Db 421 AATCTTAAATACAGATTTTGGGCTTGCAGAGGCGTTTGGTTCCTATGCGTGGTTAT 480
QY 472 ACCGATATGAGTACAGAGATGTTACGAGCTCTGAACTCTTTGCGGAGATCTCAG 531
Db 481 ACTCATGAAGTGTGTACATTTTGGGTATCGTGTCCAGAAAGTCTTTTAGTGGTGGTCGACAA 540
QY 532 G 532
Db 541 G 541

TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...894
OTHER INFORMATION:
US-09-338-125-1

Query Match 15.6%; Score 162; DB 4; Length 1476;
Best Local Similarity 55.0%; Pred. No. 1.1e-35;
Matches 362; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

QY 1 ATGGAAAGTAGTGAATAATAGCTAAGACTGGAGAGGCTCTTATGGGTTGTATTCAAA 60
DB 1 ATGGAGAACTTCCAAAGAGTGGAAAGATCGGAGAGGCGACGTACGGAGTTGTGTACAA 60
QY 61 TCCAGAAACAAACCTCTGGCAAGTAGTAGTGTGTAAATAATTTGTGAATCTGAAGAT 120
DB 61 GCCAGAACAAAGTTGACGGGAGAGGTGTGGGCTTAAGAAATCCCGCTGGACACTGAG 120
QY 121 GATCTGTGTGTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180
DB 121 ACTGAGGCTGTGCCAGTACTGCCATCCGAGAGATCTCTGCTTAAGGAGCTTAACCAT 180
QY 181 CCAATCTGTGACCTCATCGAGGTGTTACAGGAGAAAGGAAATGCAATTTAGTTT 240
DB 181 CTAATATGTCAAGCTGCTGGA-GTCATTACACAGAAATAAATCTCTACCTGGT 240
QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCCAATGAGTTGCT 297
DB 241 GAATTTCTGCCAAGATCTCAAGAAATTCATGGATGCTCTGCTCTCACTGGCATTCCT 300
QY 298 GATGAGGTGATCAAGAGCTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATA 357
DB 301 CTTCCTCCATCAAGAGCTATCTGTTCCAGCTGTCCAGGGCCCTAGCTTTCTGCCATTCT 360
QY 358 CATAACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATA 417
DB 361 CATCGGCTCTCCACGAGACTTAACCTCAGAACTCTGCTTATTACACAGAGGGGCC 420
QY 418 ATCAAGATTGTGACTTCGGGTTTGCACAAAATTCG---ATTCCAGAGAGTGCCTACACC 474
DB 421 ATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTGGAGTCCCTGTTTCGTACTACACC 480
QY 475 GATTATGTAGCTACGAGATGTTACGAGCTCTCGAATCTCTGTTGGAGATACCTCAGTAT 534
DB 481 CATGAGGTGTGACCTGTGGTACCGAGCTCTGAAATCTCTGCTGGGCTCGAATATTAT 540
QY 535 GGTTCCTCAGTCGATATATGGGCTATTGGTTGTGTTTTCGAGAGCTCTCGACAGGCCAG 594
DB 541 TCCACAGCTGTGGACATCTGGAGCTGGGCTGCATCTTTGCTGAGATGTCACCTCGCCGG 600
QY 595 CCAGTGGCTGGAATCAGATGTGACCAACTTATCTGATAATCAGAACACTAG 652
DB 601 GCCCTGTTCCCTGGAGATTCTGAGATTGACCAAGCTCTTCCGGATCTTTCGGACTCTGG 658

RESULT 7

US-09-266-225D-13

Sequence 13, Application US/09266225D

Patent No. 6573364

GENERAL INFORMATION:

APPLICANT: Nandabalan, Krishan

APPLICANT: Kingsmore, Stephen

APPLICANT: Tchernev, Velizar

TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak

TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-

TITLE OF INVENTION: Interacting Proteins

FILE REFERENCE: 15966-523

CURRENT APPLICATION NUMBER: US/09/266,225D

CURRENT FILING DATE: 1999-03-10

NUMBER OF SEQ ID NOS: 19

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 13
LENGTH: 1476
TYPE: DNA
ORGANISM: Homo sapiens
US-09-266-225D-13

Query Match 15.6%; Score 162; DB 4; Length 1476;

Best Local Similarity 55.0%; Pred. No. 1.1e-35;

Matches 362; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

QY 1 ATGGAAAGTAGTGAATAATAGCTAAGACTGGAGAGGCTCTTATGGGTTGTATTCAAA 60
DB 1 ATGGAGAACTTCCAAAGAGTGGAAAGATCGGAGAGGCGACGTACGGAGTTGTGTACAA 60
QY 61 TCCAGAAACAAACCTCTGGCAAGTAGTAGTGTGTAAATAATTTGTGAATCTGAAGAT 120
DB 61 GCCAGAACAAAGTTGACGGGAGAGGTGTGGGCTTAAGAAATCCCGCTGGACACTGAG 120
QY 121 GATCTGTGTGTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180
DB 121 ACTGAGGCTGTGCCAGTACTGCCATCCGAGAGATCTCTGCTTAAGGAGCTTAACCAT 180
QY 181 CCAATCTGTGAACTCATCGAGGTGTTACAGGAGAAAGGAAATGCAATTTAGTTT 240
DB 181 CTAATATGTCAAGCTGCTGGA-GTCATTACACAGAAATAAATCTCTACCTGGT 240
QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCCAATGAGTTGCT 297
DB 241 GAATTTCTGCCAAGATCTCAAGAAATTCATGGATGCTCTGCTCTCACTGGCATTCCT 300
QY 298 GATGAGGTGATCAAGAGCTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATA 357
DB 301 CTTCCTCCATCAAGAGCTATCTGTTCCAGCTGTCCAGGGCCCTAGCTTTCTGCCATTCT 360
QY 358 CATAACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATA 417
DB 361 CATCGGCTCTCCACGAGACTTAACCTCAGAACTCTGCTTATTACACAGAGGGGCC 420
QY 418 ATCAAGATTGTGACTTCGGGTTTGCACAAAATTCG---ATTCCAGAGAGTGCCTACACC 474
DB 421 ATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTGGAGTCCCTGTTTCGTACTACACC 480
QY 475 GATTATGTAGCTACGAGATGTTACGAGCTCTCGAATCTCTGTTGGAGATACCTCAGTAT 534
DB 481 CATGAGGTGTGACCTGTGGTACCGAGCTCTGAAATCTCTGCTGGGCTCGAATATTAT 540
QY 535 GGTTCCTCAGTCGATATATGGGCTATTGGTTGTGTTTTCGAGAGCTCTCGACAGGCCAG 594
DB 541 TCCACAGCTGTGGACATCTGGAGCTGGGCTGCATCTTTGCTGAGATGTCACCTCGCCGG 600
QY 595 CCAGTGGCTGGAATCAGATGTGACCAACTTATCTGATAATCAGAACACTAG 652
DB 601 GCCCTGTTCCCTGGAGATTCTGAGATTGACCAAGCTCTTCCGGATCTTTCGGACTCTGG 658

RESULT 8

US-09-417-197-112

Sequence 112, Application US/09417197

Patent No. 6518021

GENERAL INFORMATION:

APPLICANT: Ole THASTRUP, et al.

TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An

TITLE OF INVENTION: On A Cellular Response

FILE REFERENCE: 3759-0110P

CURRENT APPLICATION NUMBER: US/09/417,197

CURRENT FILING DATE: 1999-10-07

NUMBER OF SEQ ID NOS: 143

SOFTWARE: Patent In version 3.0

SEQ ID NO 112

LENGTH: 1635

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: CDK2-EGFP fusion
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1632)
US-09-417-197-112

Query Match 15.6%; Score 162; DB 4; Length 1635;
Best Local Similarity 55.0%; Pred. No. 1.1e-35;
Matches 362; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

```
QY 1 ATGGAAGAATATGAATAATTAGCTTAAGACTGGAGAGGGTCTTATGGGTGTGATTCAAA 60
D 1 ATGGAGAACTTCCAAAGAGTGGAAAGATCGGAGAGGGACGACGTACCGAGTGTGTACAAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAAGTAGTAGCTTAAATAATTTTGGAAATCTGAAGAT 120
D 61 GCCAGAAACAAAGTTGACGGGAGAGGTGGTGGCCCTTAAAGAAATCCGCCCTGGACACTGAG 120
QY 121 GATCCTGTTGTTAAGAAATAGCACTAGAGAAATAGCTATGTTGAAGCAATTAACAAT 180
D 121 ACTGAGGTTGTCCTGACGACTGCTCCATCCGAGAGATCTCTGCTTAAGGAGCTTAACCAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAGGAAATGCAATTTAGTTTTT 240
D 181 CCTAATATTGCAAGCTGCTGGATGTCATTCACAGAAATAAACTCTACCTGGTTTTT 240
QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCAAATGAGTTCGT 297
D 241 GAATTTCTGCACCAAGATCTCAAGAAATTCATGATGCTCTGCTCTCACTGGCATTCCT 300
QY 298 GATGGAGTATCAAAAGCGTATTATGCAACACTTCAAGCTCTTAATTTCTGTGCTATATA 357
D 301 CTTCCTCCCTCATCAAGACTATCTGTTCCAGCTCTCCAGGGCTAGCTTTCTGCCATTCT 360
QY 358 CATAACTGTATTACAGAGATATAAAACCTGAAATATTCTTAATACTAAGCAAGAAATA 417
D 362 CATCGGGTCTCCACCGAGACTTAAACCTCAGAAATCTGCTATTAAACAGAGGGGGCC 420
QY 418 ATCAAGATTGTGACTTGGGTTTGGCAAAATCTG---ATTCCAGAGATGCTTACACC 474
D 421 ATCAAGCTAGCACTTGGACTAGCCAGAGCTTTGGAGTCCCTGTCTGCTACTTACACC 480
QY 475 GATTATGACTAGAGATGGTACCGAGCTCTCGAACTCTTGTGGAGATACCTAGTAT 534
D 481 CATGAGTGTGACCTCTGGTACCGAGCTCTGAAATCTCTCGGCTCGAAATATTAT 540
QY 535 GGTCTTTCAGTCGATATATGGCTATTGGTTGTTGTTTTCAGAGCTCTTCAGAGGCCAG 594
D 541 TCCACAGCTGTGGACATCTGGAGCTGGGCTGCAATCTTTGCTGAGATGTGACTCGCCGG 600
QY 595 CCAGTGGCTCGAAATCAGATGTGGACCAACTTTATCTGATATCAGAACACTAG 652
D 601 GCCCTGTTCCTCGAGATCTGAGATTGACCACTCTCTCCGATCTTTCGGACTCTGG 658
```

RESULT 9

US-09-417-197-114

Sequence 114, Application US/09417197

Patent No. 6518021

GENERAL INFORMATION:

APPLICANT: Ole THASTRUP, et al.

TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An

TITLE OF INVENTION: On A Cellular Response

FILE REFERENCE: 3759-0110P

CURRENT APPLICATION NUMBER: US/09/417,197

CURRENT FILING DATE: 1999-10-07

NUMBER OF SEQ ID NOS: 143

SOFTWARE: PatentIn version 3.0

SEQ ID NO 114

LENGTH: 1635

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: EGFP-CDK2 fusion
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1632)
US-09-417-197-114

Query Match 15.6%; Score 162; DB 4; Length 1635;
Best Local Similarity 55.0%; Pred. No. 1.1e-35;
Matches 362; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

```
QY 1 ATGGAAGAATATGAATAATTAGCTTAAGACTGGAGAGGGTCTTATGGGTGTGATTCAAA 60
D 1 ATGGAGAACTTCCAAAGAGTGGAAAGATCGGAGAGGGACGACGTACCGAGTGTGTACAAA 798
QY 61 TGCAGAAACAAACCTCTGGACAAAGTAGTAGCTTAAATAATTTTGGAAATCTGAAGAT 120
D 61 GCCAGAAACAAAGTTGACGGGAGAGGTGGTGGCCCTTAAAGAAATCCGCCCTGGACACTGAG 858
QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAACAAT 180
D 859 ACTGAGGTTGTCCTGACGACTGCTCCATCCGAGAGATCTCTGCTTAAGGAGCTTAACCAT 918
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAGGAAATGCAATTTAGTTTTT 240
D 919 CCTAATATTGCAAGCTGCTGGATGTCATTCACAGAAATAAACTCTACCTGGTTTTT 978
QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCAAATGAGTTCGT 297
D 979 GAATTTCTGCACCAAGATCTCAAGAAATTCATGATGCTCTGCTCTCACTGGCATTCCT 1038
QY 298 GATGGAGTATCAAAAGCGTATTATGCAACACTTCAAGCTCTTAATTTCTGTGCTATATA 357
D 1039 CTTCCTCCCTCATCAAGACTATCTGTTCCAGCTCTCCAGGGCTAGCTTTCTGCCATTCT 1098
QY 358 CATAACTGTATTACAGAGATATAAAACCTGAAATATTCTTAATACTAAGCAAGAAATA 417
D 1039 CATCGGGTCTCCACCGAGACTTAAACCTCAGAAATCTGCTTTTAAACAGAGGGGGCC 1158
QY 418 ATCAAGATTGTGACTTGGGTTTGGCAAAATCTG---ATTCCAGAGATGCTTACACC 474
D 1159 ATCAAGCTAGCACTTGGACTAGCCAGAGCTTTGGAGTCCCTGTCTGCTACTTACACC 1218
QY 475 GATTATGACTAGAGATGGTACCGAGCTCTCGAACTCTTGTGGAGATACCTAGTAT 534
D 1219 CATGAGTGTGACCTCTGGTACCGAGCTCTGAAATCTCTCGGCTCGAAATATTAT 1278
QY 535 GGTCTTTCAGTCGATATATGGCTATTGGTTGTTGTTTTCAGAGCTCTTCAGAGGCCAG 594
D 1279 TCCACAGCTGTGGACATCTGGAGCTGGGCTGCAATCTTTGCTGAGATGTGACTCGCCGG 1338
QY 595 CCAGTGGCTCGAAATCAGATGTGGACCAACTTTATCTGATATCAGAACACTAG 652
D 1339 GCCCTGTTCCTCGAGATCTGAGATTGACCACTCTCTCCGATCTTTCGGACTCTGG 1396
```

RESULT 10

US-09-220-132-3

Sequence 3, Application US/09220132

Patent No. 6506607

GENERAL INFORMATION:

APPLICANT: Shyjan, Andrew W.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT

TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANC

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

CURRENT FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

```
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-220-132-3
Query Match      14.8%   Score 154.2; DB 4; Length 1050;
Best Local Similarity 54.3%; Pred. No. 1.4e-33;
Matches 359; Conservative 0; Mismatches 293; Indels 9; Gaps 2;

QY 1 ATGGAAAGATATGAAAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTAATCAA 60
DB 127 ATCGAAGATTATACCAAAATAGAGAAAATTTGGAGAGGTACCTATGGAGTTGTATAAG 186

QY 61 TCGAGAAACAAACCTTCGACAAAGTAGTACCTGTTAAAAAATTTGTGGMATCTGAAGAT 123
DB 187 GGTAGACACAAATACAGGTCAAGTGAAGTATGACCATGAAAAAATTCAGACTAGAAAGTAA 246

QY 121 GATCCTCTTTTAAAGAAAATAGCACTAAGAGAAATACGTATGTGAAGCAATTAATAACAT 180
DB 247 GAGGAAGGGGTTCTTAGTACTCGCAATTCGGGAATTTCTCTATTAAGGNACTTCGTCAT 306

QY 181 CCAATCTTTGGAACCTTCATCGAGTGTTCAGGAGAAAAAGAAATGCAATTTAGTTTTT 240
DB 307 CCAATATAGTCAGTCTCAGGATGTGCTTATGAGGATTCAGGATTCAGGTTATATCTCACTTT 366

QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGG-----AGTT 294
DB 367 GAGTTTCTTTCCATGGATCTGAAGAAATACCTTGATTTCTATCCCTCTGCTGTCAGTACATG 426

QY 295 GCTGATGGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGTCAAT 354
DB 427 GATTTCTTCACTTGTTAAGAGTTATTATACCAATCTTACAGGGGATTTGTGTTTGTCTAC 486

QY 355 ATACATTAACCTGTATTCACAGAGATATAAACCTTGAAATATTCCTAATACTAAGCAAGGA 414
DB 487 TCTAGAAGAGTCTCTTCACAGAGACTTAAACACCTCAAAATCTCTTGATGATGAAGAAGGA 546

QY 415 ATATCAAGATTTGTGACTTCGGGTTTGCACAAATCTG---ATTCCAGAGATGCTTAC 471
DB 547 ACAATTAACCTGGCTGATTTTGGCTTTCACAGAGCTTTTGGAAATACCTATCAGAGATATAT 606

QY 472 ACCGATTTATGTAGCTACGAGATGTACCGAGCTCCTSAACCTCTTGTGGGAGATACCTCAG 531
DB 607 ACATGAGGTAGTAACACTCTGTACAGATCTCCAGNAGTAATGCTGGGGTCAAGTCTGT 666

QY 532 TATGGTTCTTCAGTCSATATATGGCTATTGTTGTGTTTTCAGAGCTCTTCAGAGGC 591
DB 667 TATCAACTCCAGTTGACATTTGGATATAGGCACCATATTTGCTAACTAGCAACTAAG 726

QY 592 CAGCCACTGTGGCTGAAAAATCAGATGTGGACCAACTTTTATCTGATTAATCAAGAACTA 651
DB 727 AAACCACTTTTCCAATGGGATTTAGAAATTTGATCAACTCTTCAGGAATTTTCAGAGCTTTG 786

QY 652 G 652
DB 787 G 787
```

RESULT 11

```
US-09-620-312D-313
; Sequence 313, Application US/09420312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyun
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
```

```
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radjoe T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620.312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 313
; LENGTH: 1825
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (234)..(1235)
; US-09-620-312D-313

Query Match      12.9%   Score 134.8; DB 4; Length 1825;
Best Local Similarity 52.4%; Pred. No. 4.8e-28;
Matches 345; Conservative 0; Mismatches 307; Indels 6; Gaps 2;

QY 1 ATGNAAGATATGAAAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
DB 318 ATGGATATGTTCCAGAGGTAGAGAGATCGGAGAGGCACCTATGGGGTGGTGTACAAG 377

QY 61 TCGAAGAAACAAACCTCTCGACAAAGTAGTACCTGTAAAAAATTTGTGGAATCTGAAGAT 120
DB 378 GCCAAGAACAGGGAGACAGGGCAGCTGGTGGCCCTGAAGAGATCAGACTGGATTTGGAG 437

QY 121 GATCCTGTTTGAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAATAACAT 180
DB 438 ATGGAGGGGGTCCCAAGCACTGCCATCAGGAGATCTCGTGTCTCAAGGAACCTGAAGCAC 497

QY 181 CCAATCTTTGGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 240
DB 498 CCAACATCTGTCGACCTGCTGGACGTGTGCACAAACAGAGAGAGCTCTATCTCGTGT 557

QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAAACCCAAATGGAG---TTGCT 297
DB 559 GAGTTCTCTCAGCAGGACCTGAAGAGTACATGACTCCACCCAGGCTCAGAGCTCCCC 617

QY 298 GATGAGTGTATCAAAAGGTATTTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATA 357
DB 618 CTGCACCTCATCAAGAGCTACCTCTTCCAGCTGTGCGGGGGTGAGTTTCTGCACTCA 677

QY 359 CATAACTGTATTTCACAGAGATATAAACCTGAAATATTTCTAATAACTNAGCAAGGAATA 417
DB 678 CATCGGTCTATCCACCGAGACCTGAAGCCCCAGAACCTGCTCATCAATGATTTGGGTGCC 737

QY 418 ATCAAGATTTGTGACTTTCGGGTTTGCACAAATTTCT---GATTTCCAGGAGATGCTCATAC 474
DB 738 ATCAAGCTGGCTGACTTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 797

QY 475 GATTATGTAGCTACGAGATGGTACCGAGCTCTCTGAACTTCTTTGTGGGAGATATCTCAT 534
DB 798 CATGAGGTGGTGACACTGTGGTATCGCGCCCCCGAGATTTCTTTGGGCGAGCAAGTTCTAT 857

QY 535 GGTCTTTCAGTCTGATATATGGCTATTGTTGTGTTTTCGAGAGCTCTTCGACAGGCCAG 594
DB 858 ACCACAGCTGTGGATATCTGGAGCAATTTGGTGTGATCTTTTCAGAGATGGTGACTCGAAAA 917

QY 595 CCATGTGGCTGGAAAAATCAGATGTGACCAACTTTTATCTGATAAATCAGAACTAG 652
```


TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-464-517-37

Query Match 12.8%; Score 133.2; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 1..1e-27;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;
QY 1 ATGGAAGATGATGAAATTTAGCTAGACTGGAGAGGCTCTTATGGGTTGTATTCAA 60
DB 13 ATGCAGAAATACGAGAACTGGAAGATTTGGGAAGGCACCTACGGAACCTGTGTTCAAG 72
QY 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTTAAAGAAATTTGTGAAATCTGAAGAT 120
DB 73 GCCAAACCGGAGACTCTAGATCGTGGCTCTAAACGGGTGAGCTGSGATGACGAT 132
QY 121 GATCCTGTTTGAAGAAATAGCACTAAGAGAAATACGTAATGTTGAAGCAATTTAAACAT 180
DB 133 GATGAGGTTGTGCGAGTTCCGCCCTCCGGAGACTCTGCTACTCAAGGAGCTGAAGCAC 192
QY 181 CCAATCTGTGAACCTCATCGAGTGTTCAGGAGAAAAAGGAAATTCATTTAGTTTTT 240
DB 193 AAGAACATCGTCAGGCTTCATGACGCTCTGACAGCGACAGAGAGCTGACCTTTGGTTTTT 252
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAATACGTAATTTGAAGCAATTTAAACAT 180
DB 253 GAAATCTGTGACCGAGCTGGAAGAGTATTTTGACAGTTTGAATGTTGAGCTCGATCCT 312
QY 301 GGAGTGATCAAAAGCGTATTATGCAACACTTCGAGCTCTTAATTTCTGTCTATATCAT 360
DB 313 GAGATTGTAAGTCACTTCTCTTCCAGCTACTAAAGGGCTGGGATTTCTGTATAGCCGC 372
QY 361 AACTGTATTCACAGAGATATAAAGCTGGAAGATTTCTAATACCTAAGCAAGAAATATC 420
DB 373 AATGCTGATCACTGCTGCTGAGAGCTTTGGGATTTCCCGTCCGCTGTACTCAGCT 492
QY 421 AAGATTGTGACTTCGGGTTTGACAAA---TCTGATTCAGGAGATGGCTACACCGAT 477
DB 433 AATGCTGATTTTGGCTGGCTGGAGCTTTGGGATTTCCCGTCCGCTGTACTCAGCT 492
QY 478 TATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACCTAGTATGGT 537
DB 493 GAGGTGCTACATGTTGTTACCCGCCACCGGATGTCTCTTTGGGCGCAAGCTGTACTCC 552
QY 538 TCTTCAGTCATATATGGGCTATTGGTGTGGT---TTCAGAGCT---CTGACAGCCGAG 594
DB 553 ACGTCCATCATATGTTGTCAGCGGCTGATCTTTGACAGGCTGGCCATGCTGCGCGG 612
QY 595 CCAGTGTGGCTGGAATCAGATGTGGAACCACTTTATCTGATTAATCAGAACACTAG 652
DB 613 CTTCTTTTCCCGCAATGATGTCGATGACCAAGTTGAAGAGGATCTTCCGACTGTGG 670

RESULT 14

US-08-246-361A-37
Sequence 37, Application US/08246361A
Patent No. 598582
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246.361A
FILING DATE: 19-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963.308
FILING DATE: 16-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888.178
FILING DATE: 26-MAY-1992
APPLICATION NUMBER: US 07/701.514
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genom:c)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-246-361A-37

Query Match 12.8%; Score 133.2; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 1..1e-27;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;
QY 1 ATGGAAGATGATGAAATTTAGCTAAGACTGGAGAGGCTCTTATGGGTTGTATTCAA 60
DB 13 ATGCAGAAATACGAGAACTGGAAGATTTGGGAAGGCACCTACGGAACCTGTGTTCAAG 72
QY 61 TGCAGAAACAAACCTCTGGACAAAGTAGTAGCTGTTAAAGAAATTTGTGGAATCTGAAGAT 120
DB 73 GCCAAACCGGAGACTCATGAGATCGTGGCTCTAAAGCGGTGAGCTGGATGACGAT 132
QY 121 GATCCTGTTTGAAGAAATAGCACTAAGAGAAATACGTAATTTGAAGCAATTTAAACAT 180
DB 133 GATGAGGTTGTGCGAGTTCCGCCCTCCGGAGACTCTGCTACTCAAGGAGCTGAAGCAC 192
QY 181 CCAATCTGTGAACCTCATCGAGTGTTCAGGAGAAAAAGGAAATTCATTTAGTTTTT 240
DB 193 AAGAACATCGTCAGGCTTCATGACGCTCTGACAGCGACAGAGAGCTGACCTTTGGTTTTT 252
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAATACGTAATTTGAAGCAATTTAAACAT 300
DB 253 GAAATCTGTGACCGAGCTGGAAGAGTATTTTGACAGTTTGAATGTTGAGCTCGATCCT 312
QY 301 GGAGTGATCAAAAGCGTATTATGCAACACTTCGAGCTCTTAATTTCTGTCTATATCAT 360
DB 313 GAGATTGTAAGTCACTTCTCTTCCAGCTACTAAAGGGCTGGGATTTCTGTATAGCCGC 372
QY 361 AACTGTATTCACAGAGATATAAAGCTGGAAGATTTCTAATACCTAAGCAAGAAATATC 420
DB 373 AATGCTGATCACTGCTGCTGAGAGCTTTGGGATTTCCCGTCCGCTGTACTCAGCT 492

QY 421 AAGATTGTGACTTCGGGTTTGCACAA--TTCTGATTCAGGAGATGCTACACCGAT 477
 DB 433 AATTGGCTGATTTTGGCTGGCTGAGCCCTTTCGGATTCCCGTCCGCTGTTACTCAGCT 492
 QY 478 TATGTAGTACAGATGTGACGGAGCTCTGAACTTCTTGGGAGATACAGTAAGT 537
 DB 493 GAGTGGTCACTGTGTACCGCCACCGGATGTCTCTTGGGCAAGCTGTAATCC 552
 QY 538 TCTTCAGTCGATATATGGGTATTTGGTTGTGTTTTCAGAGCT---CTTGACAGCCAG 594
 DB 553 ACGTCCATCGACATGTGTCAGCGGCTGCATCTTTCAGAGCTGGGCAATGCTGGCGG 612
 QY 595 CCACGTGCGCTGGAAATCAGATGTGACCACTT---TCTGATAATCAGAACACTAG 652
 DB 613 CCTCTTTTCCCGCAATGATGTGATACCACTTGAAGAGGATCTTCCGACTGCTGG 670

RESULT 15
 US-08-463-772-37
 ; Sequence 37, Application US/08463772
 ; Patent No. 6066501
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463.772
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; APPLICATION NUMBER: US 07/888,178
 ; FILING DATE: 26-MAY-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/701,514
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: M11-004C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 37:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1089 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 13..888
 ;

JS-08-463-772-37
 Query Match 12.8%; Score 133.2; DB 3; Length 1089;
 Best Local Similarity 52.3%; Pred. No. 1.1e-27;
 Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

QY 1 ATGAAAGTATGAAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
 DB 13 ATGAGAAATACGAGAAATCGGAAAGATTGGGAAAGCAGCTACGGAATGTGTCAAG 72
 QY 61 TGAGAAACAAACCTCTGGAGAGTGTAGTGTGTTAAATAATTTGTGAAATCTCAAGAT 120
 DB 73 GCCAAACACCGGAGAGCTCATGAGATCGTGGCTCTAAACGGGTGAGGCTGGATCGCAT 132
 QY 121 GATCTGTTTAAAGAAATAGCATTAAAGAGAAATACGTATGTTGAAGCAATTAATAAT 180
 DB 133 GATGAGGTTGTGCGAGTTCCGCCCTCCGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192
 QY 181 CCAATCTCTGAACTCATCTGAGGTGTTTCAGAGAAAGGAAATGCAATTTAGTTT 240
 DB 193 AAGNAATCGTCAGGCTTCATGACGCTGACAGGCAAGAGCTGACTTTGTTTTT 252
 QY 241 GATATCTGTGATCATACATTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
 DB 253 GAATTTCTGTGACGAGGACCTGAAGAGATTTTTCAGAGTTGCAATGGTCACTCGATCCT 312
 QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACATTCAGCTCTTTAATTTCTGTCAATACAT 360
 DB 313 GAGATTGTAAGTCATTCTCTTCCAGCTACTAAAGGCTGGGATTTCTGTATAGCCGC 372
 QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATATC 420
 DB 373 AATGTGTACACAGGAGACCTGAAGCCCAAGAACCTGCTAATAACAGGAATGGGAGCTG 432
 QY 421 AAGATTGTGACTTCGGGTTTGCACAAA--TTCTGATTCAGGAGATGCTTACACCGAT 477
 DB 433 AATTTGGCTGATTTTGGCTGGCTCGAGGCTTGGGATTTCCGCTCGCTGTTACTCAGCT 492
 QY 478 TATGTAGCTACGAGATGGTACGAGCTCCTGAACCTTTGTGGGAGATACCTAGPATGGT 537
 DB 493 GAGGTGTCACACTGTGTGTACCCCAAGGATGCTCTTTGGGCAAGCTGTAATCC 552
 QY 538 TCTTCAGTCGATATATGGGCTATTGGTGTGTTTGTGACAGGCT---CCTGACAGGCCAG 594
 DB 553 ACGTCCATCGACATGTGGTCAGCCGCTGCATCTTTGACAGAGCTGGCAATGCTGGCGG 612
 QY 595 CCACTGTGGCTGGAAATCAGATGTGACCAACTTTATCTGATTAATCAGAACTAG 652
 DB 613 CCTCTTTTCCCGCAATGATGTGATGACCACTTGAAGAGGATCTTCCGACTGCTGG 670

Search completed: November 15, 2003, 22:31:35
 Job time : 77.9079 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 16:06:46 ; Search time 2620.85 Seconds
(without alignment)
9653.738 Million cell updates/sec

Title: US-09-671-050-9

Perfect score: 1041

Sequence: 1 agggaaagtatgaaaaatt.....aggtaacttcgcgtcaaaagt 1041

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:

2: em_esthum:

3: em_estin:

4: em_estmu:

5: em_estov:

6: em_estpl:

7: em_estro:

8: em_estc:

9: gb_est1:

10: gb_est2:

11: gb_estc:

12: gb_est3:

13: gb_est4:

14: gb_est5:

15: em_estfun:

16: em_estom:

17: em_gss_hum:

18: em_gss_inv:

19: em_gss_pln:

20: em_gss_vrt:

21: em_gss_fun:

22: em_gss_mam:

23: em_gss_mus:

24: em_gss_pro:

25: em_gss_rtd:

26: em_gss_ptg:

27: em_gss_vr:

28: gb_gsal:

29: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	480.6	46.2	695	14	CB169554 RUC603000
2	405	38.9	670	14	BY733578 BY733578
3	360.2	34.6	689	14	BY752739 BY752739
4	333.6	32.0	500	9	AI385966 ml34h09.y

5	332	31.8	1691	11	AK016781
6	316.8	30.4	938	9	AK061797
7	306.8	29.5	598	9	AV986182
8	298.4	28.7	824	13	BU221831
9	295.6	28.4	536	13	BU686325
10	289.2	27.8	597	12	BP016186
11	284.6	27.3	825	14	CB315094
12	269.8	25.9	565	14	CB430506
13	258.5	24.8	731	9	AA106692
14	255	24.5	718	9	AA286088
15	249	23.9	508	12	B15.11582
16	247.6	23.8	512	13	BX304025
17	239	23.0	485	9	AV960213
18	223.8	21.5	688	13	BM241505
19	221.8	21.3	2572	11	AK030598
20	213.4	20.5	570	12	BJ000534
21	213	20.5	1064	13	BU120363
22	210.8	20.2	627	14	BY716459
23	209.8	20.2	579	9	AM233105
24	209.4	20.1	380	13	BY075395
25	202.2	19.4	2005	11	AK090045
26	202.2	19.4	2769	11	AK045036
27	200.6	18.3	1893	11	AK050990
C 28	199	19.1	295	9	AA628659
29	199	19.1	2388	11	AK046394
30	199	19.1	2665	11	AK045356
31	197.2	18.9	471	14	CB732963
32	189.6	18.2	2013	11	AK078804
33	188.6	18.1	659	13	BM280090
34	187	18.0	673	13	BM036816
35	186.6	17.9	727	9	AI508835
36	186.2	17.9	666	14	BY733092
37	185.2	17.8	2604	11	AK052380
C 38	183.4	17.6	645	14	CA334835
39	183.4	17.6	812	13	BU377342
40	182	17.5	320	13	BY121897
C 41	180.6	17.3	774	13	BM110560
42	175.6	16.9	1535	11	BC010966
43	174.8	16.8	363	10	BB868874
44	174.8	16.8	582	9	AV672492
45	174.4	16.8	665	28	AQ953706

ALIGNMENTS

RESULT 1 CB169554 695 bp mRNA linear EST 30-JAN-2003
LOCUS RUC603000752.R1 CSEOFXN20 testes Bos taurus cDNA, mRNA sequence.
DEFINITION CB169554
ACCESSION CB169554
VERSION CB169554.1 GI:28155681
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 695)
AUTHORS Adelson,D.L. and Gill,C.A.
TITLE Bovine ESTs (Adelson and Gill)
JOURNAL Unpublished
COMMENT Contact: David L. Adelson
Animal Breeding and Genetics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
USA
Tel: 9798452616
Fax: 9798456970
Email: david.adelson@tamu.edu.

FEATURES
Location/Qualifiers
1..695
/organism="Bos taurus"


```

/mol_type="mrna"
/db_xref="taxon:9913"
/tissue_type="testes"
/clone_lib="CSEFXN20"
/notes="Organ: testes; Vector: pBluescript SK+; Site_1: Not1; Site_2: EcoRI; sequence 5' of the insert (5'-NNN...NNNinsert)
GGCAATTCGAGCTCCACCGCGTGGCGCGCGCTCGAG. Sequence 3' of the inserts: (AGAAATTCGATATCAGATATTCGATACCGTCCGACCTCGAG. normalized rd 1 library, sequenced 3' with M13R primer."
BASE COUNT 169 a 166 c 133 g 227 t
ORIGIN
Query Match 46.2%; Score 489.6; DB 14; Length 695;
Best Local Similarity 92.8%; Pred. No. 3.3e-22;
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
CY 1 ATGGAAAGTATGAAATAGCTAGACTGGAGAGGCTTTATGGSTTGTTATTCAAA 60
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
543 ATGGGAAGTATGAAATAGCTAGACTGGAGAGGCTTTATGGSTTGTTATTCAAA 484
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTTTAAAAATTTTGGAAATCTGAAGAT 120
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
483 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTTTAAAAATTTTGGAAATCTGAAGAT 424
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 GATCTGTGTTTAAAGAAATAGCATAGAGAAATACATATGTTGAAGCAATTAACAT 180
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
423 GATCTGTGTTTAAAGAAATAGCATAGAGAAATACATATGTTGAAGCAATTAACAT 364
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 CCAATCTTGTCAACTCATCGAGTGTTCAGGAGAAAGGAAATGCACTTAGTTT 240
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
363 CCAATCTTGTCAACTCATCGAGTGTTCAGGAGAAAGGAAATGCACTTAGTTT 304
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 GAATAGTGTATACACTTTAAATGAGTGTGAGAAACCCCAATGAGTGTGAT 300
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
303 GAGTACTGTGATCACACTTTAAATGAGTGTGAGAAACCCCAATGAGTGTGAT 244
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 CGAGTGTATCAAGGGTATTATGCGAAGCTTCAAGTCTTATTTCTGTCTATACAT 360
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
243 GAGTGTATCAAGGGTATTATGCGAAGCTTCAAGTCTTATTTCTGTCTATACAT 184
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 AACTGTATTCACAGAGATATAAAGCTGAAATATTTCTAATCACTAAGCAAGGAATAATC 420
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
183 AATTGTATTCACAGAGATATAAAGCTGAAATATTTCTAATCACTAAGCAAGGAATAATC 124
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 AGATTGTGACTTCGGTTTGCAGAAATTCGATTCAGGAGATGCTTACACCGATTAT 480
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 AAGATCTGTGACTTCGGTTTGCAGAAATTCGATTCAGGAGATGCTTACACCGATTAT 64
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 GTAGCTACAGATGGTACCGAGCTCTGAACTTCTGTGGGAGATACACTAGTGTGTTCT 540
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 GTCGTACAGATGGTACCGAGCTCTGAACTTCTGTGGGAGATACACTAGTGTGTTCT 4
CY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 TCA 543
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3 TCA 1

```

```

RESULT 2
BY733578
LOCUS BY733578
DEFINITION BY733578 RIKEN full-length cDNA enriched, 16 days neonate male
diencephalon Mus musculus cDNA clone G630052E12 5', mRNA sequence.
ACCESSION BY733578
VERSION BY733578.1 GI:2746705
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
1 (bases 1 to 670)
REFERENCE Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Imanaka, I., Kiyosawa, H.

```

Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojohori, T., Baicarelli, R., Hill, D.P., Bult, C., Hume, D.A., Guackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Rees, D.J., Reid, J., Ring, B.Z., Rigwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, R.D., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, Watanabe, J., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yang, S., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Azakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shiragawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Sirney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 1246685

Contact: Yoshihide Hayashizaki
 Laboratory for genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suihito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, M., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kawaji, H., Kawai, J., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Osato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers
 1..670
 /organism="Mus musculus"
 /mol_type="mrna"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G630052E12"
 /sex="male"
 /tissue_type="diencephalon"

FEATURES
 source

```

/dev stage="16 days neonate"
/clone.lib="RIKEN full-length enriched, 16 days neonate
male diencephalon"
205 a 140 c 159 g 155 t 1 others
BASE COUNT 205 a 140 c 159 g 155 t 1 others
ORIGIN
Query Match 38.9%; Score 405; DB 14; Length 670;
Best Local Similarity 85.6%; Pred. No. 1.8e-17;
Xatches 450; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1 ATGAAAGTATGAAATAGCTAGAGCTGAGAGGGCTTATGGGGTGTATCAA 60
D5 145 ATGAAAGTATGAAATAGCTAGAGCTGAGAGGGCTTATGGGGTGTATCAA 204
QY 61 TGAGAAACAAACCTCTCGAAGTAGTAGCTGTTAAATAATTTCTGGAATCTGAAGAT 120
D5 205 TGAGAAACAAACCTCTCGAAGTAGTAGCTGTTAAATAATTTCTGGAATCTGAAGAT 264
QY 121 GATCTGTTGTTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
D5 265 GATCTGTTGTTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 324
QY 181 CCAATCTTGTGAACTCATCGAGGTGTTTCAGAGAAAGCAAAATCGATGATTTT 240
D5 325 CCAATCTTGTGAACTCATCGAGGTGTTTCAGAGAAAGCAAAATCGATGATTTT 384
QY 241 GAATCTGTTGATCATACATTTTAAATGAGCTGGAAGAAACCCAAATGAGGTGCTGAT 300
D5 385 GAGTACTGTGATCACACATGTTAAACGAGCTGGAGAGAAACCCAAATGAGGTGCTGAT 444
QY 301 GAGTACTGTGATCACACATGTTAAACGAGCTGGAGAGAAACCCAAATGAGGTGCTGAT 360
D5 445 GAGTACTGTGATCACACATGTTAAACGAGCTGGAGAGAAACCCAAATGAGGTGCTGAT 504
QY 361 AACTGTATTCACAGAGATTAACACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
D5 505 AATTGTATTCACAGAGATTAACACCTGAAATATTTCTAATACTAAGCAAGGAATATC 564
QY 421 AAGATTGTGACTTCGGTTTGCACAAATCTGATTCGAGAGATGCTTACACGATAT 480
D5 565 AAGATTGTGACTTCGGTTTGCACAAATCTGATTCGAGAGATGCTTACACGATAT 624
QY 481 GTAGTACAGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATA 526
D5 625 GTTGCACACAGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATA 670

RESULT 3
BY752739 RIKEN full-length enriched, adult inner ear Mus musculus
DEFINITION
ACCESSION BY752739.1 GI:27193802
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 689)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Saito, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schorbach, C.,
Gojochori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Becken, K. W., Blake, J. A., Bratt, J. S., Brusio, V., Chochik, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lehar, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perle, G.,

```

```

pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Secou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii,
Y., Itoh, M., Kadow, I., Miyazaki, A., Sakai, K., Sasaki, R., Shibata,
K., Shingawa, A., Yasuniishi, A., Yoshino, M., Waterston, R., Shibata,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE 22354693
PUBMED 1246851
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@res.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
S., Hashizume, M., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
Ishii, Y., Itoh, M., Kadow, I., Kawai, J., Kojima, Y., Kondo, S., Konno,
H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,
Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
non-redundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Kirk W. Beisel ( Boys Town National
Research Hospital 555 North 30th Street Omaha, NE 68131 USA ) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1..689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="F930014106"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"
BASE COUNT 213 a 133 c 171 g 168 t 4 others
ORIGIN
Query Match 34.6%; Score 360.2; DB 14; Length 689;
Best Local Similarity 82.7%; Pred. No. 1.1e-14;
Matches 421; Conservative 0; Mismatches 87; Indels 1; Gaps 1;

```

```

QY 1 ATGGAAAGTAGTGAAGAAATTAGCTAGACTGAGAGGGCTCTTATGCGGTGTGATTCCAA 60
D 169 ATGGAAAGTAGTGAAGAAAGCTAGCTAGACTGAGAGGGCTCTTATGCGGTGTGATTCCAA 228
QY 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTGTTAAAGAAATTTGGAGATCTGAAGAT 120
D 229 TGCAGAAACAAATCTCTGGACAGTAGTAGCTGTGTTAAAGAAATTTGGAGATCTGAAGAT 288
QY 121 GATCCTGTGTGAGAAATAGCAATAGCAATAGCAATAGCAATAGCAATAGCAATAGCAAT 180
D 289 GATCCTGTGTGAGAAATAGCAATAGCAATAGCAATAGCAATAGCAATAGCAATAGCAAT 348
QY 181 CCAAACTCTGTGAACCTCATCGAGCTGTTCAGGAGAAAGAAAGAAATGCAATTTAGTTT 240
D 349 CCAAACTCTGTGAACCTCATCGAGCTGTTCAGGAGAAAGAAAGAAATGCAATTTAGTTT 408
QY 241 GAATAGCTGTGATACATACCTTTAAATGAGCTGTGAAGAAAGAAAGAAATGCAATTTAGTTT 300
D 409 GAGTACTGTGATACATACCTTTAAATGAGCTGTGAAGAAAGAAAGAAATGCAATTTAGTTT 468
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
D 469 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 528
QY 361 AACTGTATTCACAGAGATATAAACCTGAAATATTCATTAATTAATTAATTAATTAATTAAT 420
D 529 AACTGTATTCACAGAGATATAAACCTGAAATATTCATTAATTAATTAATTAATTAATTAAT 588
QY 421 AAGATTGTGACTTCGGGTTTGGCAAAATCTCTGATTCAGGAGATGCTACACCGATTAT 480
D 589 AAGATTGTGACTTCGGGTTTGGCAAAATCTCTGATTCAGGAGATGCTACACCGATTAT 647
QY 481 GTAGTACGAGATGATACCGAGCTCTCTGA 509
D 648 TGTGCCACCAGTGTACCGAGCGCCCGA 676

```

RESULT 4
A1385966
LOCUS
DEFINITION
m134h09.y1 Striatagene mouse testis (#937308); Mus musculus cDNA clone IMAGE:513953 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KIALRE (HUMAN); mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

A1385966.1 GI:4199429
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 500)

REFERENCE
AUTHORS
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Waller,T., Gibbons,M., Pape,D., Harvey,R., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999

TITLE
JOURNAL
COMMENT
Unpublished
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63109, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the correct orientation)
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:307801

Seq primer: -40RP from Gibco
High quality sequence stop: 339.
Location/Qualifiers

BASE COUNT 145 a 114 c 121 g 120 t
ORIGIN
Query Match
Best Local Similarity 32.0%; Score 133.6; DB 9; Length 500;
Matches 408; Conservative 0; Mismatches 84; Indels 4; Gaps 2;
QY 135 GAAATAGCACTAGAGAAATACGTATGTTGAAGCAATTAAGACATCCAAATCTCTGGAA 194
D 2 GAAATAGCACTAGAGAAATACGTATGTTGAAGCAATTAAGACATCCAAATCTCTGGAA 58
QY 195 CCTCATCGAGGTGTTCAAGAGAAAAGAAAATGCAATTTAGTTTGAATCTGTGATCA 254
D 59 CCTCATCGAGGTGTTCAAGAGAAAAGAAAATGCAATTTAGTTTGAATCTGTGATCA 118
QY 255 TACACTTTAATAGCTGGAAAGAAAACCAATGGAATTTGCTGATGGAGTGATCAAAAG 314
D 119 CACACTGTAAACGAGCTGGAGAGAAAACCAATGGAATTTGCTGATGGAGTGATCAAAAG 178
QY 315 CGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACATTAATCTGATTCACAG 374
D 179 TGTGCTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACATTAATCTGATTCACAG 238
QY 375 AGATATAAAACCTGAAATATTTCTTAATTAACCAAGGAAATTAATCAAGATTTGTGACTT 434
D 239 GGATGTAACCTGAAATATTTCTTAATTAACCAAGGAAATTAATCAAGATTTGTGACTT 298
QY 435 CGGCTTTGCACAAATCTGATTCAGAGAGATGCTACACCGATATGCTAGCTACAGATG 494
D 299 TGGATTTGCACAAATCTGATTCAGAGAGATGCTACACCGATATGCTAGCTACAGATG 358
QY 495 GTACCGAGCTCTGAACTTCTTGTGGGAGATCTCACTAGTATGGTTCCTTCAGTTCGATATG 554
D 359 GTACCGAGCTCTGAACTTCTTGTGGGAGATCTCACTAGTATGGTTCCTTCAGTTCGATATG 418
QY 555 GGCTATTGTTGTTGTTTTCGAGAGCTCTTCAGAGCGACCGACCTGTGGCTGGAAAAATC 614
D 419 GGCTATTGTTGTTGTTTTCGAGAGCTCTTCAGAGCGACCGACCTGTGGCTGGAAAAATC 477
QY 615 AGATGTGGACCAACTT 630
D 478 CGAGTGGACCACTT 493

1691 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493411017 product:cyclin-dependent kinase-like 1 (CDC2-related kinase), full insert sequence.

AK016781
AK016781.1 GI:12855701
HTC; CAP trapper.
Mus musculus (house mouse)

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```

Db 417 GAGTACTGGACCAACCGTCTTCACAGCTGATCGTATCAGAGGGGGTACCGAG 476
Qy 301 GGAGTGATCAAAAGCGTATTATGGGAAACACTTCAAGCTCTTAATTTCTGTCATATCAT 360
Db 477 CCTCTCGTGAAGAACATACTTGGCAGACACTGCGAGGCTGTAAATTTCTGCCATAAAT 536
Qy 362 AACTGTATTACAGAGATATAAATCTCAAAATATCTTAATTAATCAAGCAGGATATTC 420
Db 537 AACTGCATACAGAGAGCTGAAGCGGGAATATCTCATCACCAACAGTACGCCAAT 596
Qy 421 AAGATTGTGACTTCGGGTTTGCACAAATTCGTATTTTCCAGGAGATGCTTACACCGAT 477
Db 597 AAGCTCTGTGACTTTGGGTTTCACGGCTTCTCACTGGAACAGGTGACTACTACCGGAC 656
Qy 478 TATGTAGCTACAGAGATGTCAGGAGCTCTGAACTTCTTGTGGGAGATCTCAGTATGGT 537
Db 657 TACGTGGCCACCGGTGGTACCGCTTACCCGAGCTGCTAGTGGGAGACACGAGATGGT 716
Qy 538 TCTTCAGTCGATATATGGCTATTGGTTGTGTTTTCAGAGAGCTCCTGACAGGCCACCCA 597
Db 717 CCCCCCTGTAGATGCTGGCAATGGCTGTGTGTTTGTCTGAGTTACTGTCGGAGTSCCT 776
Qy 598 CTGTGGCTGGAAATCAGATGTGACCAACTTATCTGATAATCAGAACACTAGTAGA 656
Db 777 CTATGCGCAGGAAATCTGACGTGGATCAGCTCTACCTGATAAGGAAACCCCTGGGGGA 835

```

```

RESULT 6
AA061797 338 bp mRNA linear EST 03-FEB-1997
LOCUS m13409.r1 Stratagene mouse testis (#937308); Mus musculus cDNA
DEFINITION clone IMAGE:513953 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN
KINASE KIALRE (HUMAN);, mRNA sequence.
AA061797
AA061797.1 GI:1555656
EST.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 938)
Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubouche T.,
Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
Theising B., Wylie T., Lennon G., Soares B., Wilson R. and
Waterston R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1900
Fax: 314 286 1910
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LINDL; contact the
IMAGE Consortium (info@image.lindl.gov) for further information.
MGI:307801
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 442.
Location/Qualifiers
1. .938
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/c_lone="IMAGE:513953"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse testis (#937308)"

```

FEATURES

source

```

/note="Organ: testis; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
-5' adaptor sequence: 5' GAAATTCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTATTTTATTTT 3'"
BASE COUNT 254 a 215 c 236 g 233 t
ORIGIN
Query Match 30.4%; Score 316.8; DB 9; Length 938;
Best Local Similarity 82.9%; Pred. No. Se-12; Indels 5; Gaps 3;
Matches 398; Conservative 0; Mismatches 77;
Qy 135 GAAATACACACTAAGAGAGAAATACGATGTGTGAAGCAATTAACACATCCAATCTTGTGAA 134
Db 1 GAAATATGCCCTCGGGAAATCCGTATGCTGAAG---TTGAAACACCCAAACCTCGTGAA 57
Qy 135 CTTATCAGAGTGTTCAGGAGAAAGAAAGAAATGATCTTASTTTTGAATACCTGTGATCA 254
Db 58 CTTATCAGAGTGTTCAGGAGAAAGAAAGAAATGATCTTASTTTTGAATACCTGTGATCA 117
Qy 255 TACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGGATTCGTGATGAGTGTCAAAAG 314
Db 118 CACACTGTTAACGAGCTGGAGAGAAACCCAAACGAGGATTCGTGATGAGTGTAAAG 177
Qy 315 CGTATTATGGGAAACACTTCAAGCTCTTAATTTCTGTCTATATACATACTGTATTACAG 374
Db 178 TGTGCTATGGGAAACCTTCAAGCCCTTAACTTCTGTCAACAGCACAAATGTATTATCG 237
Qy 375 AGATATAAACTCGAAATATTTCTAATACTTAACCAAGGAATTAATCAAGATTGTGACTT 434
Db 238 GATGTAAACTGAAACATCTCTAATAACCAAGGAGATGATAAGATTGTGACTT 297
Qy 435 CGGTTTGCACAAATCTGATTCAGGAGATGCTTACACCCGATTTATGTACCTACGAGATG 494
Db 298 TGGATTGCACGAATCTTAATTCAGGAGAGCGCTCACAGACTATGTTGCCACCGTG 357
Qy 495 GTACGAGCTCTGAGCTTCTTGTGGGAGATCTCAGTATGGTTCTTCAGTCGATATG 554
Db 358 GTACCGAGCGCCCGAACTTCTCGTGGGAGACACGAAGTACGGTTCCTCTGTAGACGTGTG 417
Qy 555 GGCATTCTGTTGTGTTTTCAGAGCTCTTCAGAGCGGACGCTGTGGCTGTGAAATC 614
Db 418 GGCCTCGGCTGTGTTTTCAGAG-CTCCTGACGGGTGAG-CACTCTGCGCGGAAACCC 475
RESULT 7
AA061797 598 bp mRNA linear EST 14-MAR-2002
LOCUS AV986182 Nori Satoh unpublished cDNA library, larva Ciona
DEFINITION intestinalis cDNA clone cili4b18 5', mRNA sequence.
AA061797
AA061797.1 GI:19475050
ACCESSION AV986182
VERSION AV986182.1
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
1 (bases 1 to 598)
Sato, N., Satou, Y., Kohara, Y. and Shin-i, T.
Expressed genes in Ciona intestinalis
Unpublished
Contact: Nori Satoh
Department of Zoology
Kyoto University
Kyoto 606-8502, Japan
Sakyo-ku, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1. .598
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
FEATURES
source
1. .598

```



```

Db 482 GTGGGGACACGACATATGGCCCTCTCTGTGACGTGTGGGCATAGCTGTGCTTTGGC 541
QY 577 GAGCTCTTGTGACAGCCACCCACTGTGGCTCGAAATCAGATGTGGACCAACTTATCTG 636
DB 542 GAGCTGTCTCTCGGGGTGCCACTGTGGCCGCGCAAGTCAGACCGTGGACCACTGTACCTC 601
QY 637 ATATTCAGAACACTAGTAGA 656
DB 602 ATCCGGAGAACATTTGGGGA 621

RESULT 9
BU686325/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

BU686325 536 bp mRNA linear EST 07-OCT-2002
UI-CF-DUI-adn-i-10-0-UI-sl UI-CF-DUI Homo sapiens cDNA clone
UI-CF-DUI-adn-i-10-0-UI 3', mRNA sequence.
BU686325
BU686325.1 GI:213541120
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Gene Distribution: Researchers may obtain clones from Research
Genetics (www.regen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1..536
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-adn-i-10-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/note="Organ: Lung; Vector: pTT3-Pac (Pharmacia) with a
modified polylinker; Site1: EcoR I; Site2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pTT3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC"

```

```

BASE COUNT 148 a 104 c 97 g 187 t
ORIGIN
Query Match 28.4%; Score 295.6; DB 13; Length 536;
Best Local Similarity 98.7%; Pred. No. 1.4e-10;
Matches 298; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 740 GAATACAGGAAATTAATCCCAAGACATCAATCTTTAAAAGTAACGGGTTTTTCC 799
DB 477 GAACACTAGGAAATTAATCCCAAGACATCAATCTTTAAAAGTAACGGGTTTTTCC 418
QY 800 ATGGCATCAGTATACCTGAGCCAGACATGAACTCTTGAGGAAAGTTCTCTCAGATG 859
DB 417 ATGGCATCAGTATACCTGAGCCAGACATGAACTCTTGAGGAAAGTTCTCTCAGATG 358
QY 860 TTATCTCTGGCTCTGAACTTCTAAGAGGGTGTCTGAAGATGAATCCAGATGACAGAT 919
DB 357 TTATCTCTGGCTCTGAACTTCTAAGAGGGTGTCTGAAGATGAATCCAGATGACAGAT 298
QY 920 TAACCTGTTCCTCAACTCTCTGAGAGCTCTTCTTGAATTTCTTTCAAGAGGCCAAATTA 979
DB 297 TAACCTGTTCCTCAACTCTCTGAGAGCTCTTCTTGAATTTCTTTCAAGAGGCCAAATTA 238
QY 980 AAAGAAAGCAGCTAATGAAGGAAAGAAACAGAGAGCCCAACAGGTACTTCCGCTCAAAA 1039
DB 237 AAAGAAAGCAGCTAATGAAGGAAAGAAACAGAGAGCCCAACAGGTACTTCCGCTCAAAA 178
QY 1040 GT 1041
DB 177 GT 176

RESULT 10
BP016186
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BP016186 597 bp mRNA linear EST 15-MAR-2002
BP016186 Nori Satch unpublished cDNA library, young adult Ciona
intestinalis cDNA clone ciad64e08 5', mRNA sequence.
BP016186
BP016186.1 GI:19507663
EST.
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
1 (bases 1 to 597)
Satch,N., Satou,Y., Kohara,Y. and Shin-i,T.
Expressed genes in Ciona intestinalis
Unpublished
Contact: Nori Satch
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satcheascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..597
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciad64e08"
/tissue_type="whole animal"
/dev_stage="young adult"
/clone_lib="Nori Satch unpublished cDNA library, young
adult"
BASE COUNT 184 a 100 c 139 g 173 t 1 others
ORIGIN
Query Match 27.8%; Score 289.2; DB 12; Length 597;
Best Local Similarity 71.0%; Pred. No. 3.4e-10;
Matches 397; Conservative 0; Mismatches 159; Indels 3; Gaps 1;
QY 2 TGGAAAGCTATGAAATAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAAAT 61

```

```
Db 38 TGGAGAAATATGAGAAGATAGTAGATAGGGAAGGATCTATGGCATCTGTGTTAAAT 97
Qy 62 GCAGAAACAAACCTCTCGACAGTAGTAGCTGTAAAGAAATTTGTCGAATCTCAGAGAC 123
Db 98 GTGGAAACCGTGACTCGGGCAAAATGTGGCCATCAAAAATTTGTGCGTCTGAAGATG 157
Qy 122 ATCCCTGTTTAAAGAAATAGCACTTAAGAGAAATACGTATGTTGAAGCAATTAAGACATC 181
Db 158 ACCCCCTGATCAAGAAATCCGCTGAGGGAATCAGAAATGTTAAAGCAATTAAGACACA 217
Qy 182 CAAATCTTGGAACCTCATCAGAGTGTTCAGAGAAAGAAAGAAATGATTTAGTTTGG 241
Db 218 ATAACCTGGTGAACCTGATTGAAGTTTTCGTCGAAACCGCAATTAACATCTTGTGTTGC 277
Qy 242 AATAGCTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGATG 301
Db 278 AGTATGTCATACATACAGTTTAAATGAATTGGACAAACACATAGAGGAGTACCAGAAC 337
Qy 302 GAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTGATATACATA 361
Db 338 ATCATGTTTAAACGATTGTTTGGCAAGTTCTTCAAGCTGTTCACTTTTGGCATCAGCATA 397
Qy 362 ACTGTAATCACAGAGATATAAACCCTGAAATATTCTAACTAAGCAAGGATATATCA 421
Db 398 ATTGTATACATCGTGATGTAACACGAGAGACATTAATTAATCACTAAGCAAGGCTTATCA 457
Qy 422 AGATTGTGACTTCGGGTTTGCACAAATTTCTGA---TTCCAGGAGATGCTTACACCGATT 478
Db 458 AGTTGTGACTTTGGATTTCGAAGATTTTAACTGTTCTCTGGCGAGATTACACCGACT 517
Qy 479 ATGTAGCTACGAGATGGTACCGAGCTCTGAACTCTTGTGGGAGATCTCAGTATGGTT 538
Db 518 ATGTGGCACTAGTGGTACAGAGCACTGAGCTCTCTGTCGGGAGACACTCAGTATGGCC 577
Qy 539 CTTCACTGATATATGGGC 557
Db 578 CTNCACTGATGTGGGC 596

RESULT 11
CB315094
LOCUS
DEFINITION
AGENCOURT_115264.3 NICHD Rr Pit1 Rattus norvegicus cDNA clone
IMAGE:688377 5', mRNA sequence.
CB315094
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
CB315094.1 GI:28838974
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 bases 1 to 825
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
Tissue Procurement: John C. Marshall, M.D., Ph.D
cDNA Library Preparation: CLONTECH
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM3144 row: 1 column: 16
High quality sequence stop: 468.
Location/Qualifiers
1..825
/organism="Rattus norvegicus"
/mol_type="mRNA"
```

```
/db_xref="taxon:10116"
/clone="IMAGE:688377"
/tissue_type="Pituitary"
/lab_host="DH10B"
/clone_lib="NICHD_Rr_Pit1"
note="vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI; 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCAATTATGGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGGCGGACATG-dt(30)BN-3'
where B = A, C, G and N = A, C, G, or T). Average
insert size 1.23 kb (range 0.5-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 234 a 188 c 200 g 202 t 1 others
ORIGIN
Query Match 27.3% Score 284.6; DB 14; Length 825;
Best Local Similarity 70.9%; Pred. No. 5.5e-10;
Matches 455; Conservative 0; Mismatches 90; Indels 97; Gaps 2;
QY 361 AACTGCTTTCACAGAGATATAAAACCTGAAATATTCTAACTAAGCAAGGATATATC 420
Db 96 AGCTGATCCATCGGACGTAAACCCGAAACATCTTAATGCCAAGCAGGAGTGATA 155
QY 421 AAGATTGTGACTTCGGGTTTGGCAAAATCTGATTCAGGAGATGCTACACCGATTAT 480
Db 156 AAGATTGTGACTTCGGATTTGCACGAATCTTAATCCAGGAGATGCTACACAGACTAT 215
QY 481 CTAGCTACGAGATGTCACCGAGCTCTGAACTCTTGTGGGAGATCTCAGTATGGTTCT 540
Db 216 GTTGCACCGAGTGTACCGAGCCCTTGAACTTTGTGGAGACACGAAATACGGCTCC 275
QY 541 TCAGTCGATATATGGGCTAATGTTGTGTTTTCAGAGCTCTCTGACAGCCAGCCACTG 600
Db 276 TCTGTGGACATATGGGCTGTCGGCTGTGTTTTCGAGAGCTCTGACGGTCAGCCACTC 335
QY 601 TGGCCTGCAAAATCAGATGTGGACCACTTTATCTGATATCAACACTAGTAGAGACG 666
Db 336 TGGCAGGAAATCAGATGTGGACCACTTTATCTGATCATCAGGACATT----- 385
QY 661 GGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTCTTGAGCTCAAGTGATCCACCTGCC 720
Db 386 ----- 385
QY 721 GTAGCTCTCAAGTGTGGAAATTACAGGAAATTAATCCAGACATCAATCAATCTTT 780
Db 386 -----GGGAAAGCTGATCCCAAGGCATCAGTCTATCTTT 419
QY 781 AAAAGTAAAGGCTTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
Db 420 AAAAGTAAAGGCTTTTTCATGGCATCAGTATACCTGAGGACATGGAGACTCTT 479
QY 841 GAGGAAAGTCTTCAGATGTTTCATCTGTCGCTGTCGATCTGATCTGATG-AAGGGGTGTCGAA 899
Db 480 GAAGAAAGTCTTCAAAATGTTTCAAGCTATGCTTAAAGTTTCATGAAAGGGGTGTCGAA 539
QY 900 GATGAATCCAGATGACAGATTAACTGTTCCCAACTCTCGAGAGCTCTACTTTGATTC 959
Db 540 GATGATCTCTGATGAGGCTGACCTGTGCCAGCTGTGGACAGATGACTTACTTCTGAGTC 599
QY 960 TTTTCAAGAGGCCCAAAATTAAGAAAGCAAGCAGCTGAATGAAGG 1001
Db 600 TTTTCAAGAGGCCCAAAATTAAGAAAGCAAGCAGCTGAATGAAGG 641

RESULT 12
CB400506
LOCUS
DEFINITION
OSTf177f10_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
ACCESSION
VERSION
CB400506.1 GI:30742233
KEYWORDS
EST.
```


SOURCE
ORGANISM
 Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE
AUTHORS
 1 (bases 1 to 565)
 Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong
 C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson
 J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S.,
 Endress, G.A., Jena, S., Chevret, E., Papsotirooulos, V., Tolia, P.,
 Peacock, J., Snyder, M., Huang, R., Chance, M.R., Lee, H.,
 Doucette-Stamm, L., Hill, D.B. and Vidal, M.

TITLE
 C. elegans ORFome version 1.1: experimental verification of the
 genome annotation and resource for proteome-scale protein
 expression

JOURNAL
COMMENT
 Nat. Genet., (2003) In press
 Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739
 Email: Marc.Vidal@dfci.harvard.edu
 Sequence tag of Gateway entry clones. The primers used were
 designed on the predicted protein encoding ORF. C. elegans ORFome
 cloning project : Contact david_hill@dfci.harvard.edu or
 marc_vidal@dfci.harvard.edu
 POLYA=No.

FEATURES
source
 location/Qualifiers
 1..565
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /db_xref="taxon:6239"
 /sex="Hermaphrodite and male"
 /tissue_type="whole animal"
 /tissue="mixed stage"
 /dev_stage="AD-wrmcDNA"
 /clone_lib="AD-wrmcDNA"
 /note="The AD-wrmcDNA library was generated with poly(A)-
 RNA isolated from both hermaphrodite and male N2 worms of
 all larval stages, embryos, adults and dauers and the
 subsequent generation of cDNAs by poly(A) priming. The
 cDNAs were cloned into pPC86"
 BASE COUNT 184 a 88 c 135 g 158 t

ORIGIN
 Query Match 25.9%; Score 269.8; DB 14; Length 565;
 Best Local Similarity 67.6%; Pred.No. 5.7e-09;
 Matches 379; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

31 GGAGAGGGTCTTATGGGCTTGTTATTCAAATCGAAGAAACAAACCTCTGGACAAGTAGTA 90
 |||||
 5 GGTGAGGTTTCATATGGAGTTGTTATATATATATATATATATATATATATATATATAT 64
 |||||
 91 GCTGTTAAAGAAATTTGTGGAAATCTGAAGATGATCCTGTTGTTTGAAGAAATAGCACTAGA 150
 |||||
 65 GCAATCAAAAAGTTTGTGAAAATCTGAAGATGATCGGCATATTAAGAAAATTCGCGTACGG 124
 |||||
 151 GAATACGATGTTTGAAGCAATTAAGCAATCAAAATCTTGTGAACTTCATCGAGGTGTC 210
 |||||
 125 GAATCAGATGCTGAAGCACTGAACATCAAAATTTGGTTGGATGATGTTGAAGTGTTC 184
 |||||
 211 AGAGAAAAGGAAAATGCAATTGATTTTGAATAGTGTGATCATACACTTTTAATAGAG 270
 |||||
 185 AAACGGAACCGGAATCTCATCTCGTCTTTGAATCTTGTGATCGAATCTGCTACTTCACGAG 244
 |||||
 271 CTGGAAGAAACCAATGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 330
 |||||
 245 TTGAGAGAGATCCGATGAGTTACGATGATGATGATGATGATGATGATGATGATGATGAT 304
 |||||
 331 CTTCAAGCTCTTAATTTCTGTCATATACATAACTGTATTACAGAGATATAAAACCTGAA 390
 |||||
 305 CTAGAGCTCTTAATTTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 364
 |||||

33: AATATTCATTAATACTAAGCAAGGATATCAAGATTGTTGACTTCGGTTTCACAAATT 450
 |||||
 365 AACATTTCTTTGACACCGAATGATCAAGTGAACCTTGAGAGATTTCGATTTGTCGAATA 424
 |||||
 451 CTGATTCAGAGAGATGCTACACCGATTATGTAGTACGAGATGGTACCGAGCTCTCTGAA 510
 |||||
 425 ATAAACACGACGGAATGTATCTACTGACTATGTGCAACTCGTTGGTACCGTAGTCCAGAA 484
 |||||
 511 CTTCTTGGGAGATATCTAGTATGTTCTTCAATGATATATATATATATATATATATATATAT 570
 |||||
 485 CTTCTGTCGCGTATGTTCAATATATGGCCACCTGTAGACATTGGGCTGTAGGATGTGA 544
 |||||
 571 TTTCAGAGCTCTCTGACAGGC 591
 |||||
 545 TATCGGAGCTGTTGACTGGC 565

RESULT 13
LOCUS
DEFINITION
 um32d03.y1 Sugano mouse kidney mKia Mus musculus cDNA clone
 IMAGE:2236229 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 731)
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
 E., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter
 E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Willson, R.

REFERENCE
AUTHORS
 The WashU-NCI Mouse EST Project 1999
 Unpublished
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LIND; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1006441
 Seq primer: custom primer used
 High quality sequence stop: 433.
 Location/Qualifiers
 1..731
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:2236229"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse kidney mKia"
 /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
 (CACTGTGTG); Site_2: DraIII (CACTATGTG); 1st strand cDNA
 was primed with an oligo(dT) primer
 [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
 ligated to a DraIII adaptor [GTGGGCTACTGG], digested
 and cloned into distinct DraIII sites of the pME18S-FL3
 vector (5' site CACTGTGTG, 3' site CACTATGTG). XhoI should
 be used to isolate the cDNA insert. Size selection was
 performed to exclude fragments <1.5kb. Library
 constructed by Dr. Sumio Sugano (University of Tokyo
 Institute of Medical Science). Custom primers for
 sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end
 primer CGACTGCAGCTGAGCACA."

BASE COUNT	202 a	179 c	191 g	157 t	2 others
CRIGIN					

```
Query Match      24.8%; Score 258.6; DE 9; Length 732;
Best Local Similarity 66.9%; Pred. No. 2.5e-08;
Matches 382; Conservative 0; Mismatches 186; Indels 3; Gaps 2;
```

QY	5	AAAGATGAAAMA:TTAGCTAAGACTCGAGAGGGTCTATGSGGTGTATTTCAAATGCA	54
DB	156	AAATATGAAAANAANAATCGAAAGATTGGAGAAGGCTCTATGSGGTAGTGTTCAAAGTGA	215
QY	65	GAACAAAAACCTCTGGACACAGTAGTACGTGTAAATAAATTTGSGAACTCGGAAGATGATC	124
DB	216	GAACACGGACACCGGTGAGATCGTGGCCATCAAGAGGTTTCTGGAAACCGAAGATGACC	275
QY	125	CTGTGTTTAAAGAAAATAGCACCTAAGAGAAATACGTATGTTGAAGCAATTAACAACATCCAA	184
DB	276	CTGTGCTAAGAANAATCGCCCTCGAGANAATCCGATGCTCAAGCAACTCAAGCACCCCA	335
QY	185	ATCTGTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAGAAATGCAATTTAGTGTTTGAAAT	244
DB	336	ACCTGGTCAACCTCTCTGGAAGTCTTCCGGAGGAAGCGGAGCTTCACCTGGTGTTCGAGT	395
QY	245	ACTGTGATCATACACTTTTAAATGNGCTGGAAAGAAACCCAAATGCAGTTGCTGATGGAG	304
DB	396	ACTGGACCAACGGTGGTTCACGAGCTGGATCGGTATCAGAGGGGGTACCAGAGCCCTC	455
QY	305	TGATCAAAAGCGTATTATTGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATACATAACT	364
DB	456	TCGTGAAGAACATAACTTTGGCAGACACTGCAGGCTGTAAATTTCTGCCATAAACAATACT	515
QY	365	GTATTACACAGAGATATAAACTGAAAATATTCTAATACTAAGCAAGGAATAATCAAGA	424
DB	516	GCATACACAGACGCTGAAGCGGAAAATATTCTCATCCAAACAGTCAAGCCATTAAGC	575
QY	425	TTTGTGACTTTGGGTTTCSCAAAATTCGTATT--CCAGGAGATGCCTACACCGGATATTG	481
DB	576	TTTGTGACTTTGGGTTCCGACGGTCTCTCACTGACCANGTGCTACTACACAGACTACN	635
QY	482	TAGCTACGAGATGGTACCGAGCTCTCGAATCTTCTTGTGGGAGATACCTCAGTATGGTCTT	541
DB	636	GTGGCCACCCGTGGTACTGCTCAACCGAGCTGCTAGTTGGAGACACGCGATATTGGTCCC	695
QY	542	CAGTCGATATATGGGCTATTGGTGTGTTTT	572
DB	696	CTGTAGATGCTCGGGCAATTCGGCTGTGTGTT	726

RESULT 14
AA286988
LOCUS
DEFINITION
718 bp "RNA linear EST 09-APR-1997
AA286089
VC3B05.r1 Barstead MELB1 Mus musculus cDNA clone IMAGE:776337 5'
similar to GD:X66356 SERINE/THREONINE-PROTEIN KINASE KK1ALRE (HUYAN
); mRNA sequence.

ACCESSION	AA286088
VERSION	AA286088.1
KEYWORDS	G:1932198
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

mus. muscivorus
Cranium

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
1 (bases 1 to 718)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steproe, M., Tan, F., Underwood, K., Moore, P.,
Theisling, B., Wylie, T., Lennon, G., Soares, S., Wilson, R. and
Waterston, R.

REFERENCE
AUTHORS

TITLE	JOURNAL	COMMENT
WashU-HHMI Mouse EST Project	Unpublished	Contact: Marita M/Mouse EST Project
WashU-HHMI Mouse EST Project	Washington University School of Medicine	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

```

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG:469193
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 424.
Location/Qualifiers
    1. 718
       /organism="Mus musculus"
       /mol_type="mRNA"
       /strain="BALB/c"
       /db_xref="taxon:10090"
       /clone="IMAGE:776337"
       /sex="mixed"
       /tissue_type="Kidney"
       /dev_stage="6 weeks"
       /lab_host="DH10B"
       /clone_lib="Barstead MPLRB1"
       /note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site1: EcoRI; Site2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTCAAGATGGAGCGCCGCTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CATGATTCGGTACC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."

```

BASE COUNT	199 a	171 c	186 g	162 t	Library constructed by EOD, Paris, France.
ORIGIN					
Query Match	24.5%	Score 255;	DB 9;	Length 718;	
Best Local Similarity	67.9%	Pred. No. 4.2e-08;			
Matches 416;	Conservative	0;	Mismatches 190;	Indels 7;	Gaps 4;
QY	1	ATGGAAGAATGATGA	AAATTAGCTA	AGACTGAGAGAGGGTCTTATGGGTTGTATTCAAA	60
DB	109	ATGGAAGAATGATGA	AAATATGGA	AAATATGGAAGAGATTGGAAGAGGCTCTCTATGGGTTAGTGTCTCAAG	168
QY	61	TGCAGAAACAAACCTCTG	GACAAAGTAGT	AGCTGTAAAAAATTTGCGAATCTGGAAGAT	120
DB	159	TGCAGAACAGGACAC	CGGTCAGATCTG	TGGCCATCAAGAGTTCTTGGAAACCGAAGAT	228
QY	121	GATCCTCTGTTTAA	GAATAATAGCATA	AGAGAAATACGTAATGTTGAAGCAATATAAAACAT	180
DB	229	GACCTCTGATTA	AGAAATATCGCCCTTC	GAGAAATATCCGATGCTCAAGCAACTCAAGCAC	288
QY	281	CCAATCTCTGAACTCAT	CGAGGTGTTTCAGGACAA	AAAGGAAATATGCAATTTAGTTTTT	240
DB	289	CCAACTCTGTCAACTCT	CTGGAATCTTCCGAGG	-AAGCGAGGCTTCACCTGGTGTTC	347
QY	241	GAATCTGTGATCAT	ACATCTTTTAAATGAGCT	TGGAAAGAAACCCAAATGGAGTTGCTGTAT	300
DB	348	GAGTACTCGCAACCA	CCGTCGCTCACGAGCT	GGATCGGTA-TCAGAGGGGGGTACAGAG	406
QY	301	GGAGTGTATCAAAAG	CGTATATATGGCAACACTTCA	AGCTCTTATTTCTGTCATATACAT	360
DB	437	CCTCTCGTGAAGA	CACTAACTTGGCAGACACT	TGCAGGCTGTAAATTTCTGCCATAAAACAT	466
QY	361	AACTGTATTACAGAG	ATATAAACCTGAAAAATATTTCTAA	TAACTTAAGCAAGGAATATATC	420
DB	467	AACTGCATACAGAG	ACGTGAAGCGGAAATATTTCTATT	ACCACAAGTCAGCCATT	526
QY	421	AAGATTTGTGACTTC	GGGTTTGA-----CAAATTTCT	GAATCCAGAGATGCTTACACCGA	476
DB	527	AAGCTCTGTCACTTT	GGGGTTGCAACGGCTTCTCACT	GGAGCCAGCTGACTACTACACGGA	586
QY	477	TTATGTAGCTACGAG	ATGTTACCGAGCTCTCTGAACTTCT	CTGTGGGAGATCTCAGTAGTG	536
DB	587	CTACGTGGCCACCC	GGTGGTACCCGTCAACCGAGCT	TGCTAGTGGGAGACACGCAAGTAGTG	646
QY	537	TTCTTCAGTCGATAT	TGGGCTATTGGTTGTGTTTTTC	GCAGAGCTCTCTGACAGGCCACGCC	596

Db 647 TCCCCCTGAGTGTCTGGGC-AATGGCTGTGTGTTGCTGAGTACTGTCCGAGTGCC 705
 QY 597 ACTGTGGCTGGG 609
 Db 706 TCTATGGCAAGG 718

RESULT 15

BI511582 508 bp mRNA linear EST 08-APR-2002
 LOCUS BB160006A20C12.5 Bee Brain Normalized Library, BB16 Apis mellifera
 DEFINITION cDNA clone BB160006A20C12 5', mRNA sequence.

ACCESSION BI511582
 VERSION
 KEYWORDS

SOURCE EST. BI511582.1 GI:15361956

ORGANISM Apis mellifera (honeybee);

Apis mellifera

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;

Apidae; Apis.

1 (bases 1 to 508)

REFERENCE Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L.,

Pardinas,J., Robertson,H.M., Soares,B. and Robinson,G.E.

Annotated expressed sequence tags and cDNA microarrays for studies

of brain and behavior in the honey bee

Genome Res. 12 (4), 555-566 (2002)

21929762

11932240

Contact: Gene E. Robinson

Department of Entomology

University of Illinois

505 S. Goodwin Ave., Urbana, IL 61801, USA

Tel: 217 265 0309

Fax: 217 244 3499

Email: gerober@life.uiuc.edu

This research was funded by the University of Illinois Critical

Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation

Award in Functional Genomics to G.E. Robinson and an NSF

Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.

PCR Primers

FORWARD: TAATACGACTCACTATAGGG

BACKWARD: ATTACCCCTCACTAAG

Plate: BB160006A20 row: C column: 12

Seq primer: ACCGGATAACAATTTCACACAGGA

High quality sequence stop: 508.

Location/Qualifiers

1..508

/organism="Apis mellifera"

/mol_type="mRNA"

/strain="mixed strains of European bees, predominantly

A.m. ligustica"

/db_xrefs="taxon:7460"

/clone="BB160006A20C12"

/sex="female"

/tissue_type="brain"

/dev_stage="adult worker honey bee"

/lab_host="DH:03"

/clone_lib="Bee Brain Normalized Library, BB16"

/note="Organ: Brain; Vector: pT73-pac; Site:1: EcoRI;

Site:2: NotI; The BB16 library was contributed by the

Soares laboratory and it was constructed and normalized

as described by Bonaldo, M.F., Lennon, G. and Soares,

M.B. (1996). Genome Research 6(9): 791-806. RNA was

prepared from dissected brains of adult worker bees of

various ages and various behavioral groups."

BASE COUNT 163 a 93 c 118 g 134 t

ORIGIN

Query Match 23.9%; Score 249; DB 12; Length 508;

Best Local Similarity 68.3%; Pzed.No. 1.2e-07;

Matches 345; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 110 AATCTGAAGATGATCCTGTTTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGC 169

Db 4 AAACAGAGATGATCCTCTCTTATACGAAATATGGCTACGTGAAATAGACTTCTTAAGA 63
 QY 170 AATTAACATCCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGAAATGC 229
 Db 64 ATCTAAACATCCAAATCTGGTCAATCTCTGGAGGTCTTTCAGACGAAACGGAATTAC 123
 QY 230 AATTAGTTTTTGATATGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATG 269
 Db 124 ACCTAGTCTTTGAATATTCGAGATATAGTTTAAATGAGATGGAGATATCCGAGTG 183
 QY 290 GAGTTGCTGATGAGTGAATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTTAATTTCT 349
 Db 184 GGTGTCGGGAATCACTACAAACAACTCACATGCGCAATACTGCAGGCGCATCGCATATT 243
 QY 350 GTCATATACATACTGATTTCACAGAGATATAAACTGAAATATTTCTAATACTAAGC 409
 Db 244 GCCATCGATTGGGTGTGTTCATAGAGATGTAAGCGCGGAAATATTTTAATCACAGCGG 303
 QY 410 AAGGAATAATCAAGATTGTGACTTCGGGTTTGCAAAATTTCTGATTCCAGGAGATGCCT 469
 Db 304 ATGGCGTGTAATTTGTGTGACTTCGGTTTCGCTCGAATGCTCAGTCTGGTGAAATTT 363
 QY 470 ACACCGATTATGTAGCTACAGAGATGGTACCGAGCTCTGAACTTCTTGTGGAGATACTC 529
 Db 364 ACACAGATATGTTGCAACTAGATGGTACAGAGCACTGAACTATTGGTTGGAGATACTC 423
 QY 530 AGTATGGTTCTTCAGTGCATATATGGGCTATTGGTTGTGTTTTTGAGAGCTCCTGCAG 589
 Db 424 AATACGGTACACCACTCGATGTATGGGCAATCGGTTGTGTTTTTCGAGAACTAATACGTG 483
 QY 590 GCCAGCCACTGTGGCTGGGAAATC 614
 Db 484 GAGAAGCACTATGGCCAGGAAATC 508

Search completed: November 15, 2003, 20:27:44

Job time : 2627.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:36:36 ; Search time 327.082 Seconds
(without alignments)
8591.481 Million cell updates/sec

Title: US-09-671-050-9

Perfect score: 1041

Sequence: 1 atggaagaagtatgaanaatt.....aggtacttcgcgcacaaagt 1041

Scoring table: IDENTITY NUC

Gapop 10_C, Gapext 1.0

Searched: 2552756 seqs, 134919017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1041	100.0	1041	22	Human kinase cDNA
2	1035	98.5	1068	22	Human kinase cDNA
3	839	80.6	945	22	Human kinase cDNA
4	823	79.1	972	22	Human kinase cDNA
5	821	78.9	1790	24	Human kinase cDNA
6	762.8	73.3	1678	25	Human kinase cDNA
7	757	72.7	1083	22	Human kinase cDNA
8	709	68.1	1819	22	Human kinase cDNA

3	644	61.9	882	24	AA164248	Human kinase 14257
10	543.2	52.2	1281	25	ABX34679	Human mdt cDNA SE
11	537	51.6	911	22	AAF44669	Novel protein cDNA
12	454.4	43.7	561	22	AAD03812	Human kinase cDNA
13	454.4	43.7	594	22	AAD03815	Human kinase cDNA
14	453.2	43.5	2615	22	AAF44670	Novel protein cDNA
15	395	37.9	1086	25	ABZ77165	Human protein cDNA
16	343.8	33.0	2944	23	ABQ93433	Human cDNA SEQ ID
17	343.8	33.0	2944	23	ABV24584	Human prostate exp
18	251.8	24.2	1698	21	AAZ29746	Rabbit KIAMRE kin
19	251.8	24.2	3080	21	AAZ29745	Rabbit KIAMRE kin
20	242.2	23.3	1179	23	ABL12603	Drosophila melanog
21	202.6	19.5	1513	21	AAZ51208	Human lost in leuk
22	202.6	19.5	1773	22	AAZ51207	Polynucleotide seq
23	191	18.3	1667	21	AAZ51206	Rat lost in leukae
24	191	18.3	1738	21	AAZ51206	Rat lost in leukae
25	186.6	17.9	3297	23	ABL12602	Drosophila melanog
26	177.4	17.0	1158	23	AAZ70329	DNA encoding novel
27	175.2	16.8	1866	22	AAF44672	Novel protein cDNA
28	163.8	15.7	903	20	AAZ37835	Pneumocystis carin
29	162	15.6	1476	21	AAZ37835	Cyclin-dependent k
30	162	15.6	1635	19	AAV71073	CDK2-green fluore
31	162	15.6	1635	19	AAV71074	Green fluorescent
32	161.6	15.5	1024	25	ABZ83372	Toxicologically re
33	161.6	15.5	1297	24	ABQ60787	Human 10Kshel cDN
34	160.4	15.4	2213	24	ABK84073	Human cDNA differe
35	160	15.4	1322	21	AAF15889	Human prostate can
36	154.2	14.8	1050	24	ABT11085	Human breast cance
37	154.2	14.8	1050	24	ABL64774	Lung cancer relate
38	154.2	14.8	1050	24	ABL64774	Lung cancer relate
39	154.2	14.8	1050	24	ABL65855	Lung cancer relate
40	154.2	14.8	1050	24	ABL65855	Kidney cancer rela
41	154.2	14.8	1050	25	ACC50096	Breast cancer asso
42	154.2	14.8	1050	25	AAD52783	Human CDC2 DNA, H
43	154.2	14.8	1235	25	ABX76248	Lung cancer-associ
44	154.2	14.8	1825	24	AAZ94983	Human DNA sequence
45	154.2	14.8	1897	25	ACA03966	cDNA downregulated

ALIGNMENTS

RESULT 1
AAD03816
ID AAD03816 standard; cDNA; 1041 BP.
XX AAD03816;
AC AAD03816;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human kinase cDNA #5.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW Biological disorder; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FI CDS 1..1041
FT /*tag= a
FT /*product= "Human kinase #5"
FT /*note= "The coding region does not include stop codon."
FT /*partial
XX
FN WO200123579-A1.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX WPI; 2001-266166/27.
 DR P-PSDB; AAE00494.
 XX
 PT New isolated human kinase polynucleotide useful for generating
 PT antibodies, as reagents in diagnostic assays and for screening for
 PT compounds useful for treating mental, biological or medical diseases
 XX
 PS Claim 1; Page 31-32; 38pp; English.
 XX
 CC The present sequence is a cDNA encoding novel human protein
 CC (NHP) known as human kinase. The human kinases share structural
 CC similarity with animal kinases, more particularly serine or
 CC threonine protein kinases. Human kinase cDNA is useful for the
 CC detection of mutant human kinase for the diagnosis of disease,
 CC and also as a therapeutic. It is useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic
 CC manifestations perturbing the normal function of NHP in the
 CC body. The NHP nucleotide sequences are useful for generation of
 CC antibodies as reagents in diagnostic assays, for the
 CC identification of other cellular gene products related to human
 CC kinases, and as reagents in assays for screening compounds that
 CC are useful for treating mental, biological or medical disorders.
 CC NHP oligonucleotides are used as probes. The labelled NHP probes
 CC are useful for screening human genomic library for identifying
 CC polymorphisms and as primers in amplification assays to detect
 CC mutations within the exons, introns and splice sites that can
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct
 CC encoding NHP products are used to genetically engineer cells
 CC in vivo that functions as bioreactors in the body delivering a
 CC continuous supply of NHP to the body. Nucleotide constructs
 CC encoding functional NHPs are used in gene therapy for the
 CC modulation of NHP expression.
 XX
 SQ Sequence 1041 BP; 344 A; 190 C; 228 G; 279 T; 0 other;

XX Query Match 100.0%; Score 1041; DB 22; Length 1041;
 XX Best Local Similarity 100.0%; Pred. No. 1.5e-268;
 XX Matches 1041; Conservative C; Mismatches C; Indels 0; Gaps 0;

QY 1 ATGGAAAGTATGAATAATAGCTAAGACTGAGAGGGCTCTATGGGTTGATTCAAA 60
 DB 1 ATGGAAAGTATGAATAATAGCTAAGACTGAGAGGGCTCTATGGGTTGATTCAAA 60
 QY 61 TCGAGAAACAAACCTCTGGACAAGTAGTACGTGTTAAATAATTTGGAAATCTGAAGAT 120
 DB 61 TCGAGAAACAAACCTCTGGACAAGTAGTACGTGTTAAATAATTTGGAAATCTGAAGAT 120
 QY 121 GATCTGTGTTAAGNAATAGCCTAGAGNAATACGTATGTTGAGCAATTAAACAT 180
 DB 121 GATCTGTGTTAAGNAATAGCCTAGAGNAATACGTATGTTGAGCAATTAAACAT 180
 QY 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCAATTTAGTTT 240
 DB 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCAATTTAGTTT 240
 QY 241 GAATAGTGTGATCATACATTTTAAATAGCTGGAGAAAGAAACCCCAATGGAGTGTGAT 300
 DB 241 GAATAGTGTGATCATACATTTTAAATAGCTGGAGAAAGAAACCCCAATGGAGTGTGAT 300
 QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
 DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
 QY 361 AACTGTATTTCAGAGATATAAAACCTGAAATATTTCTAATTAACCTAAGCAAGGAATAATC 420
 DB 361 AACTGTATTTCAGAGATATAAAACCTGAAATATTTCTAATTAACCTAAGCAAGGAATAATC 420
 QY 421 AAGATTGTGACTTCGGGTTTCGCAAAATTCGTATCCAGGAGATGCTACACCGATTAT 480
 DB 421 AAGATTGTGACTTCGGGTTTCGCAAAATTCGTATCCAGGAGATGCTACACCGATTAT 480

481 GTAGCTACGAGATGGTACCAGCTCCTGAACCTCTTGTGGGAGATACTCAGTATGGTTCT 540
 DB 481 GTAGCTACGAGATGGTACCAGCTCCTGAACCTCTTGTGGGAGATACTCAGTATGGTTCT 540
 QY 541 TCAGTCCGATATATGGGCTATTGGTTGTGTTTTTGGCAGAGCTCCTGACAGGCCAGCCACTG 600
 DB 541 TCAGTCCGATATATGGGCTATTGGTTGTGTTTTTGGCAGAGCTCCTGACAGGCCAGCCACTG 600
 QY 601 TGGCTCGAAATAATCAGATGTGGACCAACTTTATCTGATATCAGAAACATAGTAGAGAGG 660
 DB 601 TGGCTCGAAATAATCAGATGTGGACCAACTTTATCTGATATCAGAAACATAGTAGAGAGG 660
 QY 661 GGGTTTCGCCATGTTGACACAGGCTGGTCTCGAACTCTTGACGTCAAGTATCCACCTGCC 720
 DB 661 GGGTTTCGCCATGTTGACACAGGCTGGTCTCGAACTCTTGACGTCAAGTATCCACCTGCC 720
 QY 721 GTAGCCTCTCAAGTGTCTGGAATTTACAGGAAATTTAATCCCAAGACATCAATCAATCTTT 780
 DB 721 GTAGCCTCTCAAGTGTCTGGAATTTACAGGAAATTTAATCCCAAGACATCAATCAATCTTT 780
 QY 781 AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACACATGGAAATCTTT 840
 DB 781 AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACACATGGAAATCTTT 840
 QY 841 GAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAATTCATGAAGGGGTCTCTGAAG 900
 DB 841 GAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAATTCATGAAGGGGTCTCTGAAG 900
 QY 901 ATGAATCCAGATGACAGATTAACCTGTTCCCAACTCTCGAGAGCTCCTACTTTGATTCT 960
 DB 901 ATGAATCCAGATGACAGATTAACCTGTTCCCAACTCTCGAGAGCTCCTACTTTGATTCT 960
 QY 961 TTTCAAGAGGCCCAAAATTTAAAGAAAGACGTAATTAAGAAAGAAACAGAGACGCCAA 1020
 DB 961 TTTCAAGAGGCCCAAAATTTAAAGAAAGACGTAATTAAGAAAGAAACAGAGACGCCAA 1020
 QY 1021 CAGTACTTCGGCTCAAAAGT 1041
 DB 1021 CAGTACTTCGGCTCAAAAGT 1041

RESULT 2
 AAD03823
 ID AAD03823 standard; cDNA; 1068 BP.
 XX AC AAD03813;
 XX DT 19-JUN-2001 (first entry)
 XX Human kinase cDNA #2.
 XX Human; kinase; gene therapy; bioreactor; mental disorder;
 KW biologic.al disorder; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 CDS 1..1068
 FT /*tag= a
 FT /product= "Human kinase #2"
 FT /note= "The coding region does not include stop codon"
 FT /partial
 XX WO200123579-A1.
 XX 05-APR-2001.
 XX 27-SEP-2000; 2000WO-US26621.
 XX 28-SEP-1999; 99US-0156511.
 XX (LEXI-) LEXICON GENETICS INC.

PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands A*;
 XX WPI: 2001-266166/27.
 DR P-PSDB; AAE00491.
 XX
 PT New isolated human kinase polynucleotide useful for generating
 PT antibodies, as reagents in diagnostic assays and for screening for
 PT compounds useful for treating mental, biological or medical diseases
 PS Disclosure; Page 28; 38pp; English.
 XX
 CC The present sequence is a cDNA encoding novel human protein
 CC (NHP) known as human kinase. The human kinases share structural
 CC similarity with animal kinases, more particularly serine or
 CC threonine protein kinases. Human kinase cDNA is useful for the
 CC detection of mutant human kinase for the diagnosis of disease,
 CC and also as a therapeutic. It is useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic
 CC manifestations perturbing the normal function of NHP in the
 CC body. The NHP nucleotide sequences are useful for generation of
 CC antibodies, as reagents in diagnostic assays, for the
 CC identification of other cellular gene products related to human
 CC kinases, and as reagents in assays for screening compounds that
 CC are useful for treating mental, biological or medical disorders.
 CC NHP oligonucleotides are used as probes. The labelled NHP probes
 CC are useful for screening human genomic library for identifying
 CC polymorphisms and as primers in amplification assays to detect
 CC mutations within the exons, introns and splice sites that can
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct
 CC encoding NHP products are used to genetically engineer cells
 CC in vivo that functions as bioreactors in the body delivering a
 CC continuous supply of NHP to the body. Nucleotide constructs
 CC encoding functional NHPs are used in gene therapy for the
 CC modulation of NHP expression.
 XX
 SQ Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;
 Query Match 98.5%; Score 1025; DB 22; Length 1068;
 Best Local Similarity 100.0%; Pred. No. 2.9e-264;
 Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGGAAGATGATAAAGATAGCTAGAGCTGAGAGAGGGTCTTATGGGTGTGATTCAA 60
 1 ATGGAAGATGATAAAGATAGCTAGAGCTGAGAGAGGGTCTTATGGGTGTGATTCAA 60
 61 TCGAGAAACAAACCTCTGCAAGTAGTAGCTGTAAAGAAATTTGTGGAATCTGAGAT 120
 61 TCGAGAAACAAACCTCTGCAAGTAGTAGCTGTAAAGAAATTTGTGGAATCTGAGAT 120
 121 GATCCTGTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTAAGCAATTAAGCAT 180
 121 GATCCTGTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTAAGCAATTAAGCAT 180
 181 CCAATCTTGTGAACTCATCGAGGTGTTCAAGAGAAAGGAAATGCAATTTAGTTT 240
 181 CCAATCTTGTGAACTCATCGAGGTGTTCAAGAGAAAGGAAATGCAATTTAGTTT 240
 241 GAATACGTGTGATCAGACATCTTTAATAGCTGGAGAGAAACCAATGGAGTGTCTGAT 300
 241 GAATACGTGTGATCAGACATCTTTAATAGCTGGAGAGAAACCAATGGAGTGTCTGAT 300
 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
 361 AACTGTATTACAGAGATATAAAGCTGAAATATTTCTAATTAAGCAAGGAATATATC 420
 361 AACTGTATTACAGAGATATAAAGCTGAAATATTTCTAATTAAGCAAGGAATATATC 420
 421 AAGATTGTGACATTCGGGTTTGACAAATTTCTGATTCAGAGAGATGCCCTACACGATTTAT 480
 421 AAGATTGTGACATTCGGGTTTGACAAATTTCTGATTCAGAGAGATGCCCTACACGATTTAT 480

QY 481 GTAGCTACGAGATGGTACCGAGCTCTCGAACTTTCTTGGGAGATATCAGTATGGTCT 540
 DB 481 GTAGCTACGAGATGGTACCGAGCTCTCGAACTTTCTTGGGAGATATCAGTATGGTCT 540
 QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTGGCAGAGCTCTCGACAGGCGACCACTG 600
 DB 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTGGCAGAGCTCTCGACAGGCGACCACTG 600
 QY 601 TGCGCTGGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTAGTAGAGACG 660
 DB 601 TGCGCTGGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTAGTAGAGACG 660
 QY 661 GGGTTTCGGCATGTTTGACCGAGCTGGTCTCGAACTTTGACGTCAGAGTCAAGTGTATCCACCTGCC 720
 DB 661 GGGTTTCGGCATGTTTGACCGAGCTGGTCTCGAACTTTGACGTCAGAGTCAAGTGTATCCACCTGCC 720
 QY 721 GTAGCTCTCAAGTGTGGGATTTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
 DB 721 GTAGCTCTCAAGTGTGGGATTTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
 QY 781 AAAAGTAAAGCGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATCGAAACTCTT 840
 DB 781 AAAAGTAAAGCGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATCGAAACTCTT 840
 QY 841 GAGGAAAGTTCTCAGATGTTTCATCTGCTGGTCTGAACTTCATGAGGGGTGTCTGAAG 900
 DB 841 GAGGAAAGTTCTCAGATGTTTCATCTGCTGGTCTGAACTTCATGAGGGGTGTCTGAAG 900
 QY 901 ATGAATCCAGATGACAGATTAACCTGTTTCCCAACTCTCGAGAGCTCTCTTGTATCT 960
 DB 901 ATGAATCCAGATGACAGATTAACCTGTTTCCCAACTCTCGAGAGCTCTCTTGTATCT 960
 QY 961 TTTCAAGAGGCCCAATTAAGAAAGAAAGCACTAATGAAGGAGAAACAGAGAGCGCAA 1020
 DB 961 TTTCAAGAGGCCCAATTAAGAAAGAAAGCACTAATGAAGGAGAAACAGAGAGCGCAA 1020
 QY 1021 CAGGT 1025
 DB 1021 CAGGT 1025

RESULT 3
 AAD03817
 ID AAD03817 standard; cDNA; 945 BP.
 XX
 AC AAD03817;
 XX
 DT 19-JUN-2001 (first entry)
 XX
 DE Human kinase cDNA #6.
 XX
 KW Human; kinase; gene therapy; bioreactor; mental disorder;
 KW biological disorder; ss.
 XX
 CS Homo sapiens.
 PH Key Location/Qualifiers
 FT CDS 1..945
 FT /product= "Human kinase #6"
 FT /note= "The coding region does not include stop codon"
 XX
 PN W0200123579-A1.
 XX
 PD 05-APR-2001.
 XX
 PF 27-SEP-2000; 2000WO-US26621.
 XX
 PR 28-SEP-1999; 99US-0156511.
 XX
 PA (LEXI-) LEXICON GENETICS INC.

XX Doncho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX WPI: 2001-266166/27.
 DR P-PSDB: AA500492.
 XX New isolated human kinase polynucleotide useful for generating
 PT antibodies, as reagents in diagnostic assays and for screening for
 PT compounds useful for treating mental, biological or medical diseases
 XX
 PS Disclosure: Page 29-30; 38pp: English.
 XX
 CC The present sequence is a cDNA encoding novel human protein
 CC (NHP) known as human kinase. The human kinases share structural
 CC similarity with animal kinases, more particularly serine or
 CC threonine protein kinases. Human kinase cDNA is useful for the
 CC detection of mutant human kinase for the diagnosis of disease,
 CC and also as a therapeutic. It is useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic
 CC manifestations perturbing the normal function of NHP in the
 CC body. The NHP nucleotide sequences are useful for generation of
 CC antibodies, as reagents in diagnostic assays, for the
 CC identification of other cellular gene products related to human
 CC kinases, and as reagents in assays for screening compounds that
 CC are useful for treating mental, biological or medical disorders.
 CC NHP oligonucleotides are used as probes. The labelled NHP probes
 CC are useful for screening human genomic library for identifying
 CC polymorphisms and as primers in amplification assays to detect
 CC mutations within the exons, introns and splice sites that can
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct
 CC encoding NHP products are used to genetically engineer cells
 CC in vivo that functions as bioreactors in the body delivering a
 CC continuous supply of NHP to the body. Nucleotide constructs
 CC encoding functional NHPs are used in gene therapy for the
 CC modulation of NHP expression.
 XX
 SQ Sequence 972 BP; 323 A; 172 G; 214 G; 263 T; 0 other;

Query Match 79.1%; Score 823; DB 22; Length 972;

Best Local Similarity 90.6%; Pred. No. 3.3e-210;

Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 ATGGAAAGTATGAAATATAGCTAAGACTGAGAGGGTCTATGGGTTGTATCCAAA 60
 DB 1 ATGGAAAGTATGAAATATAGCTAAGACTGAGAGGGTCTATGGGTTGTATCCAAA 60
 QY 61 TCGAGAAACAAACCTCTCGCAAGTAGTGTGTAATAAATTTGTGGAATCTGAAGAT 120
 DB 61 TCGAGAAACAAACCTCTCGCAAGTAGTGTGTAATAAATTTGTGGAATCTGAAGAT 120
 QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAGCAATTAAACAT 180
 DB 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAGCAATTAAACAT 180
 QY 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGAGAAAGAAATGCAATTTAGTTT 240
 DB 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGAGAAAGAAATGCAATTTAGTTT 240
 QY 241 GAATCTGTGATCATACCTTTAATGAGCTGGAGAAACCCMAATGGAGTTGCTGAT 300
 DB 241 GAATCTGTGATCATACCTTTAATGAGCTGGAGAAACCCMAATGGAGTTGCTGAT 300
 QY 301 GGAGTATCAAAAGCGTATTTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
 DB 301 GGAGTATCAAAAGCGTATTTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
 QY 361 AACTGTATTACAGAGATATAAACTGGAANAATTTCTAATACTAAGCAAGGAATATC 420
 DB 361 AACTGTATTACAGAGATATAAACTGGAANAATTTCTAATACTAAGCAAGGAATATC 420
 QY 421 AAGATTGTGACTTCGGGTTTCCAGAAATTTGATTCAGAGAGATGCGCTACACCGATTAT 480
 DB 421 AAGATTGTGACTTCGGGTTTCCAGAAATTTGATTCAGAGAGATGCGCTACACCGATTAT 480

QY 481 GTAGCTACGAGATGCTACCGAGCTCCGAACTTCTGTGGAGATCTCAGTATGGTTCT 540
 DB 481 GTAGCTACGAGATGCTACCGAGCTCCGAACTTCTGTGGAGATCTCAGTATGGTTCT 540
 QY 541 TCAGTCGATATATGGCTATTGGTTGTGTTTTCGAGAGCTCCCTGACAGGCCAGCCACTG 600
 DB 541 TCAGTCGATATATGGCTATTGGTTGTGTTTTCGAGAGCTCCCTGACAGGCCAGCCACTG 600
 QY 601 TGGCTCGAAATCAGATGTCGACCACTTTATCTGATATCAGAACACTAGTAGAGACG 660
 DB 601 TGGCTCGAAATCAGATGTCGACCACTTTATCTGATATCAGAACACTAGTAGAGACG 660
 QY 661 GGGTTTCGCCATGTTGACCAAGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
 DB 661 GGGTTTCGCCATGTTGACCAAGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
 QY 721 GTAGCTCTCAAGTGTGGAAATTAACAGGAATAATTAATCCAGACATCAATCAATCTTT 780
 DB 721 GTAGCTCTCAAGTGTGGAAATTAACAGGAATAATTAATCCAGACATCAATCAATCTTT 780
 QY 781 AAAAGTAAACGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
 DB 781 AAAAGTAAACGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
 QY 841 GAGGAAAGTTCTCAGATGTTTCATCTGCTGCTGAACTTCATGAAGGGGTGTTCTGAAG 900
 DB 841 GAGGAAAGTTCTCAGATGTTTCATCTGCTGCTGAACTTCATGAAGGGGTGTTCTGAAG 900
 QY 901 ATGAATCCAGATGACAGATTAACCTGTTCCCACTCTCTGGAGAGCTCCTACTTTGATTCT 960
 DB 901 ATGAATCCAGATGACAGATTAACCTGTTCCCACTCTCTGGAGAGCTCCTACTTTGATTCT 960
 QY 961 TTTCAAGAGGCCCAATTTAAAGAAAGAACACGTAATGAAGGAAGAAAACAGAAAGCGCAA 1020
 DB 961 TTTCAAGAGGCCCAATTTAAAGAAAGAACACGTAATGAAGGAAGAAAACAGAAAGCGCAA 1020
 QY 1021 CAGGT 1025
 DB 925 CAGGT 929
 RESULT 5
 AAD30557
 ID AAD30557 standard; cDNA; 1790 BP.
 XX
 AC AAD30557;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Humar kinase polypeptide (PKIN-10) cDNA.
 XX
 KW Human: kinase polypeptide; PKIN-10; gene therapy; Addison's disease;
 KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
 KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hyperextension;
 KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
 KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
 KW cholestasis; cardiac; cardiovascular disorder; Niemann-Pick's disease;
 KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
 KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
 KW hypotensive; anti-HIV; enzyme; ss.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT CDS 239..1267
 FT /tag= a
 FT /product= "Human PKIN-10"
 XX
 PN WO200208399-A2.
 XX
 PD 31-JAN-2002.
 XX

20-JUL-2001; 2001WO-US23092.

21-JUL-2000; 2000US-220038P.

28-JUL-2000; 2000US-222112P.

04-AUG-2000; 2000US-222831P.

11-AUG-2000; 2000US-224729P.

(INCYTE GENOMICS INC.

(THORNTON M.

Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Wallia NK;

Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;

Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lai P;

Ding L, Yao MS, Elliott VS, Recipon SA, Kearney L, Lu DAM;

Greenwald SR, Tang YF, Xu Y, Walsh RT, Gietzen KJ, Yang J;

Hillman JL;

WPI: 2002-206083/26.

P-PSDB; AAE19152.

New human kinase polypeptide, useful in diagnosis, prevention and

treatment of cancer, immune disorder, growth and developmental

disorder, cardiovascular disorder and lipid disorder

Claim 5; Page 183; 196pp; English.

The present invention relates to an isolated human kinase polypeptide

(PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is

useful for diagnosing, treating and preventing cancer (e.g., leukemia,

lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency

disease, AIDS), Addison's disease, allergy, anemia, asthma, Crohn's

disease, rheumatoid arthritis), a growth and developmental disorder (e.g.

buritis, cirrhosis, hepatitis), psoriasis, Cushing's syndrome), a

cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial

infarction), and a lipid disorder (e.g., fatty liver, cholestasis,

Gaucher's disease, Niemann-Pick's disease) PKIN is useful in a number of

drug screening techniques and to analyse the proteome of a tissue or cell

type. PKIN is useful for creating knockin humanised animals or transgenic

animals to model human diseases, in somatic or germline gene therapy, to

generate a transcript image of a tissue or cell type, for detecting

differences in the chromosomal location due to translocation, inversion,

etc., among normal, carrier or affected individuals, and as hybridisation

probes for mapping naturally occurring genomic sequences. PKIN is useful

in southern or northern analysis, dot blot or other membrane-based

technologies, in PCR technologies, in dipstick, pin, multiformat enzyme

linked immunosorbent (ELISA)-like assays and in microarrays utilising

fluids or tissues from patients to detect altered PKIN expression. The

present sequence is human PKIN-10 cDNA.

Sequence 1790 BP; 581 A; 328 C; 366 G; 315 T; 0 other;

Query Match 78.9%; Score 821; DB 24; Length 1790;

Best Local Similarity 90.6%; Pred.No. 1.4e-209;

Matches 927; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

1 ATGGAAAGTATGAGAAATTTAGCTAAGACTGGAGAGCGTCTTATGGGTTGTATCAAA 60

239 ATGGAAAGTATGAGAAATTTAGCTAAGACTGGAGAGCGTCTTATGGGTTGTATCAAA 298

61 TGCAGAAACAAACCTCTCGGCAAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120

299 TGCAGAAACAAACCTCTCGGCAAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 358

121 GATCTCTGTTTGAAGAAATAGCCTTAAGAGAAATACGTATGTTGAGCAATTAACAT 180

359 GATCTCTGTTTGAAGAAATAGCCTTAAGAGAAATACGTATGTTGAGCAATTAACAT 418

181 CCAATCTTTGTGAACCTCATCGAGGTGTTTCAGAGAAAGAAATGCAATTTAGTTT 240

419 CCAATCTTTGTGAACCTCATCGAGGTGTTTCAGAGAAAGAAATGCAATTTAGTTT 478

241 GATATCTGTATCATACATTTTAAATGAGCTGGAGAAACCAACCAATGGAGTCTGAT 300

Db	479	GAACTCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGAGTGTGCTGAT	538
Qy	301	GSAGTGATCAAAACGCTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT	360
Db	539	GSAGTGATCAAAACGCTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT	598
Qy	361	AACCTGATTACACAGAGATATAAACCTGAAATATTCTAATAACTAAGCAAGAAATAATC	420
Db	599	AACCTGATTACACAGAGATATAAACCTGAAATATTCTAATAACTAAGCAAGAAATAATC	658
Qy	421	AAGATTCTGCTACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATCCCTACACCGATTAT	480
Db	659	AAGATTCTGCTACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATCCCTACACCGATTAT	718
Qy	481	GTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATACTCAGTATGTTCT	540
Db	719	GTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATACTCAGTATGTTCT	778
Qy	541	TCAGTCCGATATATGGGCTATTGGTTGTTTTTGCAGAGCTCTGACAGCGCCAGCCACTG	600
Db	779	TCAGTCCGATATATGGGCTATTGGTTGTTTTTGCAGAGCTCTGACAGCGCCAGCCACTG	838
Qy	601	TGGCTTGGAAATCAGATGTGGAGCACTTTTATCTGATATCAGAACACTAGTAGAGAGCG	660
Db	839	TGGCTTGGAAATCAGATGTGGAGCACTTTTATCTGATATCAGAACACT-----	888
Qy	661	GGGTTTGGCCATGTTGACAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC	720
Db	889	-----	888
Qy	721	GTAGCTCTCAAGTGGCTGGAATTAAGAGAAATTAATCCAGAGACATCAATCAATCTTT	780
Db	889	-----AGGAAATTAATCCAGAGACATCAATCAATCTTT	922
Qy	781	AAAGTACAGGGTCTTCCATGGCATCAGTATACCTCGAGCCAGAGACATGGAATCTTT	840
Db	923	AAAGTACAGGGTCTTCCATGGCATCAGTATACCTCGAGCCAGAGACATGGAATCTTT	982
Qy	841	GAGGAAAGTTCTCAGATGTTCTGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	983	GAGGAAAGTTCTCAGATGTTCTGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG	1042
Qy	901	ATGATCCAGATGACAGATTAACCTGTTTCCCACTCTCGAGAGCTCTTCTTCTTCTTCT	960
Db	1043	ATGATCCAGATGACAGATTAACCTGTTTCCCACTCTCGAGAGCTCTTCTTCTTCTTCT	1102
Qy	961	TTTCAAGAGGGCCCAATTTAAAGAAAGACAGTAAATGAAGAAAGAAAGAAAGAAAG	1020
Db	1103	TTTCAAGAGGGCCCAATTTAAAGAAAGACAGTAAATGAAGAAAGAAAGAAAGAAAG	1162
Qy	1021	CAG 1023	
Db	1163	CAG 1165	
RESULT 6			
ABX08936			
ID	ABX08936	standard; cDNA; 1678 BP.	
AC	ABX08936;		
XX			
DT	21-JAN-2003	(first entry)	
XX			
DE		cDNA encoding human DITHP protein #6.	
XX			
KW		Human; ss; Gene; diagnostic and therapeutic; DITHP; cancer;	
KW		arteriosclerosis atherosclerosis; psoriasis; primary thrombocytopenia;	
KW		autoimmune; inflammatory; anaemia; asthma; autoimmune thyroiditis;	
KW		Crohn's disease; diabetes mellitus; glomerulonephritis; gout; stroke;	
KW		multiple sclerosis; rheumatoid arthritis; uveitis; AIDS; allergy;	
KW		acquired immunodeficiency disease; neurological disorder; epilepsy;	
KW		Alzheimer's disease; dementia; mental retardation; gastrointestinal;	
KW		Parkinson's disease; ulcer; cirrhosis; reproductive; infertility;	

D5 1055 GCCAACAG +062

RESULT 7

AA06725
ID AA06725 standard; cDNA; 1093 BP.

XX AC
XX AA06725;

XX 12-SEP-2001 (first entry)

XX DB

XX Polynucleotide sequence encoding human protein kinase #25.

XX KW

XX Human; protein kinase; PK; STK; cancer; cardiovascular disease;

XX KW metabolic disorder; immune related disease; neurological disorder;

XX KW neurodegenerative disorder; inflammatory disorder; infectious disease;

XX KW reproductive disorder; gene therapy; ss.

XX OS

XX Homo sapiens.

XX PN

XX WO200138503-A2.

XX PD

XX 31-MAY-2001.

XX PF

XX 22-NOV-2000; 2000WO-US32085.

XX PR

XX 24-NOV-1999; 99US-0:67482.

XX PA

XX (SUGEN) SUGEN INC.

XX PI

XX Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;

XX PI Flanagan P, Ciary D;

XX XX

XX WPI; 2001-343950/36.

XX DR

XX P-PSDB; RA003525.

XX PT

XX Nucleic acids encoding human kinase polypeptides, useful for preventing

XX PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and

XX PT neuronal-associated diseases, and microbial infections -

XX XX

XX Example 1; Figure 1; 433pp; English.

XX PS

XX AAS06701-AA06757 encode for novel human protein kinases #1-57. The

XX CC novel protein kinases have been identified as members of the tyrosine

XX CC or serine/threonine kinase (PTK and STK) families. The polynucleotides

XX CC encoding protein kinases and the polypeptides may be used in the

XX CC prevention, diagnosis and treatment of diseases associated with

XX CC cancers (especially cancers of hematopoietic origin); cardiovascular

XX CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),

XX CC immune related diseases (e.g. rheumatoid arthritis), neurological

XX CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.

XX CC parkinson's disease), inflammatory disorders (e.g. asthma), infectious

XX CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).

XX CC Additionally, polynucleotides encoding protein kinases may be

XX CC used for gene therapy and as DNA probes in diagnostic assays.

XX CC The protein kinase polypeptides may be used as antigens in the production

XX CC of antibodies against the protein kinases and in assays to identify

XX CC modulators of protein kinase expression and activity.

XX XX

XX Sequence 1093 BP; 366 A; 199 C; 225 G; 293 T; 0 other;

XX SQ

Query Match 72.7%; Score 757; DB 22; Length 1083;

Best Local Similarity 86.1%; Pred. No. 1.6e-192;

Matches 927; Conservative 0; Mismatches 0; Indels 150; Gaps 2;

QY 1 ATGGAAAGTATGAAATTAAGCTAGGAGGAGGCTCTATGGGGTGTATTCAA 60

DB 1 ATGGAAAGTATGAAATTAAGCTAGGAGGAGGCTCTATGGGGTGTATTCAA 60

QY 61 TGCAGAACAAACCTCTGGCAAGTACTAGCTCTTAAATAATTTGGGAATCTGAAGAT 120

DB 61 TGCAGAACAAACCTCTGGCAAGTACTAGCTCTTAAATAATTTGGGAATCTGAAGAT 120

QY

DB

QY

DB

RESULT 8

AA03818

ID AAD03818 standard; cDNA; 1819 BP.

XX XX

AC AAD03818;

XX XX

DT 19-JUN-2001 (first entry)

121 CATCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
 121 GATCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
 181 CCAAATCTTGTGAACCTCATCGAGGTTT CAGGAGAAAAAGGAAAAATGCATTAGTTTTT 240
 181 CCAAATCTTGTGAACCTCATCGAGGTTT CAGGAGAAAAAGGAAAAATGCATTAGTTTTT 240
 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAGAAACCCCAATGAGTTGCTGAT 300
 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAGAAACCCCAATGAGTTGCTGAT 300
 301 GGAGTGTCAAAAACCGTATTATGCGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360
 301 GGAGTGTCAAAAACCGTATTATGCGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360
 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTTAATACCTAAGCAAGGAATATC 420
 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTTAATACCTAAGCAAGGAATATC 420
 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGA----- 454
 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGA----- 454
 455 -----TTCCAGGAGATGCTTACACCGATTATGTAGCT 486
 455 -----TTCCAGGAGATGCTTACACCGATTATGTAGCT 486
 481 TTACCTGATCTGATTGATGCACTTTGCAGATTTCCAGGAGATGCTTACACCGATTATGTAGCT 540
 481 TTACCTGATCTGATTGATGCACTTTGCAGATTTCCAGGAGATGCTTACACCGATTATGTAGCT 540
 487 ACGAGATGGTACCGAGCTCTTGAACTTTCTGCGAGATACCTCAGTATGTTCTTCAGTC 546
 487 ACGAGATGGTACCGAGCTCTTGAACTTTCTGCGAGATACCTCAGTATGTTCTTCAGTC 546
 541 ACGAGATGGTACCGAGCTCTTGAACTTTCTGCGAGATACCTCAGTATGTTCTTCAGTC 600
 541 ACGAGATGGTACCGAGCTCTTGAACTTTCTGCGAGATACCTCAGTATGTTCTTCAGTC 600
 547 GATATATGGGCTATTGTTGTTGTTTTCAGAGCTCTCTGACGCGCCAGCCACTGTGGCCT 606
 547 GATATATGGGCTATTGTTGTTGTTTTCAGAGCTCTCTGACGCGCCAGCCACTGTGGCCT 606
 601 GATATATGGGCTATTGTTGTTGTTTTCAGAGCTCTCTGACGCGCCAGCCACTGTGGCCT 660
 601 GATATATGGGCTATTGTTGTTGTTTTCAGAGCTCTCTGACGCGCCAGCCACTGTGGCCT 660
 607 GGAATAATCAGATGGGACCACTTTATCTGATATCAGAACACTAGTAGAGAGCGGGTTT 666
 607 GGAATAATCAGATGGGACCACTTTATCTGATATCAGAACACTAGTAGAGAGCGGGTTT 666
 661 GGAATAATCAGATGGGACCACTTTATCTGATATCAGAACACT----- 704
 661 GGAATAATCAGATGGGACCACTTTATCTGATATCAGAACACT----- 704
 667 CGCCATGTTGACAGGCTGGTCTCGAACTCTTGACGTCAGAGTATCCACCTGCGGTAGCC 726
 667 CGCCATGTTGACAGGCTGGTCTCGAACTCTTGACGTCAGAGTATCCACCTGCGGTAGCC 726
 705 ----- 704
 705 ----- 704
 727 TCTCAAGTGTGGAATTAAGCAAGATCAATCAATCAATCAATCAATCAATCAATCAAT 786
 727 TCTCAAGTGTGGAATTAAGCAAGATCAATCAATCAATCAATCAATCAATCAATCAAT 786
 705 -----AGGAAATTAATCCCAAGACATCAATCAATCAATCAATCAATCAAT 744
 705 -----AGGAAATTAATCCCAAGACATCAATCAATCAATCAATCAATCAATCAAT 744
 787 AACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAACCTCTTTGAGGAA 846
 787 AACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAACCTCTTTGAGGAA 846
 745 AACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAACCTCTTTGAGGAA 804
 745 AACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAACCTCTTTGAGGAA 804
 847 AAGTTCTCAGATGTTCTGCTGGCTCTGAACTTCAATGAGGGGTCTCTGAGAGATGAAT 906
 847 AAGTTCTCAGATGTTCTGCTGGCTCTGAACTTCAATGAGGGGTCTCTGAGAGATGAAT 906
 805 AAGTTCTCAGATGTTCTGCTGGCTCTGAACTTCAATGAGGGGTCTCTGAGAGATGAAT 864
 805 AAGTTCTCAGATGTTCTGCTGGCTCTGAACTTCAATGAGGGGTCTCTGAGAGATGAAT 864
 907 CCAGATGACATTAACCTGTTCCCAACTCTCTGAGAGCTCCTACTTTGATTCTTTTCAA 966
 907 CCAGATGACATTAACCTGTTCCCAACTCTCTGAGAGCTCCTACTTTGATTCTTTTCAA 966
 865 CCAGATGACATTAACCTGTTCCCAACTCTCTGAGAGCTCCTACTTTGATTCTTTTCAA 924
 865 CCAGATGACATTAACCTGTTCCCAACTCTCTGAGAGCTCCTACTTTGATTCTTTTCAA 924
 967 GAGGCCCAAAATTAAGAAAGACGCTAATGAAGGAAAGAAACAGACAGCCCAACAG 1023
 967 GAGGCCCAAAATTAAGAAAGACGCTAATGAAGGAAAGAAACAGACAGCCCAACAG 1023
 925 GAGGCCCAAAATTAAGAAAGACGCTAATGAAGGAAAGAAACAGACAGCCCAACAG 981
 925 GAGGCCCAAAATTAAGAAAGACGCTAATGAAGGAAAGAAACAGACAGCCCAACAG 981

XX DE Human kinase cDNA #7.
 XX DE 5'UTR 1..137
 XX FT /*tag= a
 XX FT CDS 138..701
 XX FT /*tag= b
 XX FT /*product= "Human kinase #1"
 XX FT 702..1819
 XX FT 3'UTR
 XX FT /*tag= c
 XX FT misc_difference 1684
 XX FT /*tag= d
 XX FT /*note= "polymorphism site; given as r in the sequence"
 XX PN WO200123579-A1.
 XX PD 05-APR-2001.
 XX PP 27-SEP-2000; 2006NO-US2662..
 XX PR 28-SEP-1999; 99US-0156511.
 XX PA {LEXI-} LEXICON GENETICS INC.
 XX PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX DR WPI: 2001-266166/27.
 XX DR P-PSDB; AAE00490.
 XX PT New isolated human kinase polynucleotide useful for generating
 XX PT antibodies, as reagents in diagnostic assays and for screening for
 XX PT compounds useful for treating mental, biological or medical diseases -
 XX PS Disclosure; Page 34-35; 38pp; English.
 XX CC The present sequence encodes a novel human protein (NHP) which
 XX CC is a human kinase. A polymorphism was identified in the 3' UTR
 XX CC of the present sequence. The human kinases share structural
 XX CC similarity with animal kinases, more particularly serine or
 XX CC threonine protein kinases. Human kinase cDNA is useful for the
 XX CC detection of mutant human kinase for the diagnosis of disease,
 XX CC and also as a therapeutic. It is useful for screening drugs
 XX CC effective in the treatment of symptomatic or phenotypic
 XX CC manifestations perturbing the normal function of NHP in the
 XX CC body. The NHP nucleotide sequences are useful for generation of
 XX CC antibodies, as reagents in diagnostic assays, for the
 XX CC identification of other cellular gene products related to human
 XX CC kinases, and as reagents in assays for screening compounds that
 XX CC are useful for treating mental, biological or medical disorders.
 XX CC NHP oligonucleotides are used as probes. The labelled NHP probes
 XX CC are useful for screening human genomic library for identifying
 XX CC polymorphisms and as primers in amplification assays to detect
 XX CC mutations within the exons, introns and splice sites that can
 XX CC be used in diagnostics and pharmacogenomics. Nucleotide construct
 XX CC encoding NHP products are used to genetically engineer cells
 XX CC in vivo that functions as bioreactors in the body delivering a
 XX CC continuous supply of NHP to the body. Nucleotide constructs
 XX CC encoding functional NHPs are used in gene therapy for the
 XX CC modulation of NHP expression.
 XX SQ Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;
 XX Query Match 68.1%; Score 709; DB 22; Length 1819;
 XX Best Local Similarity 77.0%; Pred. No. 1.4e-179;
 XX Matches 1025; Conservative 0; Mismatches 0; Indels 306; Gaps 1;
 XX 1 ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATCAA 60

Db 138 ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATCAA 197
 Qy 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTCAAGAT 120
 Db 198 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTCAAGAT 257
 Qy 121 GATCCTGTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAATAACAT 180
 Db 258 GATCCTGTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAATAACAT 317
 Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGAAAATGATTTAGTTTTT 240
 Db 315 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGAAAATGATTTAGTTTTT 377
 Qy 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAGAAACCCAAATCGGAGTTCTCAT 300
 Db 378 GAATCTGTGATCATACACTTTTAAATGAGCTGGAGAAACCCAAATCGGAGTTCTCAT 437
 Qy 301 GGAGTGATCAAAAGCGTAATTATGGCAACACCTCAAGCTCTTAATTTCTGTCTATACAT 360
 Db 438 GGAGTGATCAAAAGCGTAATTATGGCAACACCTCAAGCTCTTAATTTCTGTCTATACAT 497
 Qy 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAAATATC 420
 Db 498 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAAATATC 557
 Qy 421 AAGATTGTGACTTCGGGTTTGCACAAAATCTG----- 453
 Db 558 AAGATTGTGACTTCGGGTTTGCACAAAATCTGAGTTGGACTTCATCTTCTCTGGTCCC 617
 Qy 454 ----- 453
 Db 618 TCCTTGATTGGCTTAATAGTTCACCTTCTGAATTTCTTTTCTGCCAATTCAGAGATTTT 677
 Qy 454 ----- 453
 Db 678 CTCTGGCTTGGATCCATGTCACACAGTGTTCACCATGGGGCCAGGCTCATCTCGA 737
 Qy 454 ----- 453
 Db 738 ACTTCTGGCTCAAGTAGTCCTTCCACCTCGSCCTCCCAAGTGCTGGATTGCAAGTGTG 797
 Qy 454 ----- 453
 Db 798 AGCCACCGTGCCAGCCAGATTTTCAAACAATAACTACTGAGAGCTCACAAGA'TGTTTT 857
 Qy 454 -----ATTCCAGGAGATGCCCTACACC 474
 Db 858 TTAGTGGGAACACAATTTTGGAAACAAATTTTGGAGAACCGATTCAGGAGATGCCCTACACC 917
 Qy 475 GATTATGTAGCTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGATCACTCAGTAT 534
 Db 918 GATTATGTAGCTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGATCACTCAGTAT 977
 Qy 535 GGTCTTTCAGTCGATATATGGCTATTGGTGTGTTTTTGCAGAGCTCCTCAGAGGCCAG 594
 Db 978 GGTCTTTCAGTCGATATATGGCTATTGGTGTGTTTTTGCAGAGCTCCTCAGAGGCCAG 1037
 Qy 595 CCATGTGGCTTGGAAAATCAGATGTGACCAACTTTATCTGATAATCAGAACACTAGTA 654
 Db 1038 CCATGTGGCTTGGAAAATCAGATGTGACCAACTTTATCTGATAATCAGAACACTAGTA 1097
 Qy 655 GAGACGGGGTTTCGCATGTTGACACAGCTGGTCTCGAACTCTTGACGCTCAAGTGATCCA 714
 Db 1098 GAGACGGGGTTTCGCATGTTGACACAGCTGGTCTCGAACTCTTGACGCTCAAGTGATCCA 1157
 Qy 715 CTGCGCTAGCCTCTCAAAGTGTGGAATTAAGAGAAAATTAATCCCAAGACATCAATCA 774
 Db 1158 CTGCGCTAGCCTCTCAAAGTGTGGAATTAAGAGAAAATTAATCCCAAGACATCAATCA 1217
 Qy 775 ATCTTTAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAA 834

Db 1218 ATCTTTAAAGTAACGGGTTTTCATGGCATCATATACCTGAGCCAGAGACATCGAA 1277
Qy 835 ACTCTTGAGGAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAGGGGTCT 894
Db 1278 ACTCTTGAGGAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAGGGGTCT 1337
Qy 995 CTGAAGATGAATCCAGATGACAGATTAACTGTGTTCCCACTCTCTGGAGAGCTCCTACTTT 954
Db 1338 CTGAAGATGAATCCAGATGACAGATTAACTGTGTTCCCACTCTCTGGAGAGCTCCTACTTT 1397
Qy 955 GATTCTTTCAAGAGGCCCAATTTAAAGAAACACAGTATGAGGAAAGAAACAGAGA 1014
Db 1398 GATTCTTTCAAGAGGCCCAATTTAAAGAAACACAGTATGAGGAAAGAAACAGAGA 1457
Qy 1015 CGCCAAACAGGT 1025
Db 1458 CGCCAAACAGGT 1468
RESULT 9
AA164248
ID AA164248 standard; cDNA; 882 BP.
XX
AC AA164248;
XX
DT 08-MAR-2002 (first entry);
XX
DE Human kinase 14257 cDNA.
XX
KW Protein kinase; enzyme; cytostatic; osteopathic; hepatotropic;
KW antidiabetic; neuroprotective; antiarthritic; dermatological;
KW immunosuppressive; antiinflammatory; antithyroid; antipsoriatic;
KW ophthalmological; antiallergic; antiasthmatic; antiatherosclerotic;
KW hypotensive; vasotonic; antiarrhythmic; virucide; anorectic;
KW metabolic; immunomodulator; analgesic; cellular proliferative disorder;
KW cancer; acute lymphoblastic leukaemia; Hodgkin's disease;
KW bone metabolism disorder; osteoporosis; immune system disorder;
KW inflammatory; diabetes mellitus; osteoarthritis; asthma;
KW cardiovascular disorder; hypertension; coronary artery disease;
KW endothelial cell disorder; psoriasis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..126
FT CDS 129..815
FT /tag= a
FT /tag= b
FT /product= "Protein kinase 14275"
FT 3'UTR 816..882
FT /tag= c
XX
PN WO200179488-A2.
XX
XX 25-OCT-2001.
XX
XX 13-APR-2001; 2001WO-US12188.
XX
XX 13-APR-2000; 2000US-196910P.
XX
XX (MIL-) MILLENNIUM PHARM INC.
XX
XX Kapeller-libermann R;
XX
XX WPI; 2002-034355/04.
XX P-PSDB; AAG78547.
XX
XX New 14257 polypeptides (protein kinases); useful as diagnostic targets
XX PT and therapeutic agents for controlling cellular proliferative and/or
XX PT differentiative disorders, bone disorders, immune disorders and
XX PT cardiovascular disorders -
XX
PS Claim 2; Fig 1A; 98pp; English.

XX The invention relates to an isolated 14257 polypeptide and nucleic
XX acid encoding it. The 14257 protein is a protein kinase that acts as a
XX modulating agent in regulating a variety of cellular processes,
XX including cell proliferation, differentiation, growth and division.
XX The activity of the protein of the invention may be described as;
XX cytostatic; osteopathic; hepatotropic; antidiabetic; neuroprotective;
XX antiarthritic; dermatological; immunosuppressive; antiinflammatory;
XX antithyroid; antipsoriatic; ophthalmological; antiallergic;
XX antiasthmatic; antiatherosclerotic; hypotensive; vasotonic;
XX antiarrhythmic; virucide; anorectic; metabolic; immunomodulator and
XX analgesic. The protein of the invention may act as a novel diagnostic
XX target or therapeutic agent controlling certain disorders, for example
XX kinase-associated or other 14257-associated disorders. These may include
XX cellular proliferative disorders such as cancers e.g. acute lymphoblastic
XX leukemia or Hodgkin's disease. Other disorders include bone metabolism
XX disorders such as osteoporosis, disorders of the immune system, e.g.
XX inflammatory, diabetes mellitus, osteoarthritis and asthma. Proteins of
XX the invention may also be of use as therapeutic agents in cardiovascular
XX disorders such as hypertension and coronary artery disease, and some
XX endothelial cell disorders, including psoriasis. The current
XX sequence represents a human kinase 14257 cDNA.
XX
SQ Sequence 882 BP; 288 A; 150 C; 198 G; 245 T; 1 other;

Query Match 61.9%; Score 644; DB 24; Length 882;
Best Local Similarity 99.2%; Pred. No. 2.6e-162;
Matches 647; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ATGGAAGATGTAAGAAATAGCTAAGACTGGAGAGGGCTTATGGGTGTATTCAAA 60
Db 129 ATGGAAGATGTAAGAAATAGCTAAGACTGGAGAGGGCTTATGGGTGTATTCAAA 188
Qy 61 TGCAGAAACAAACCTCTGGAACAGTAGTAGCTTTAAAAAATTTGGAAATCTGAAGAT 120
Db 199 TGCAGAAACAAACCTCTGGAACAGTAGTAGCTTTAAAAAATTTGGAAATCTGAAGAT 248
Qy 121 GATCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
Db 249 GATCCTATGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 308
Qy 181 CCNAATCTTGTGAACCTCATCGAGGTGTTGAGGAGAAAAGGAAAATGCATTTAGTTT 240
Db 309 CCNAATCTTGTGAACCTCATCGAGGTGTTGAGGAGAAAAGGAAAATGCATTTAGTTT 368
Qy 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGAAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 369 GAATACCTGTGATCATACACTTTTAAATGAGCTGAAAAGAAACCCAAATGGAGTTGCTGAT 428
Qy 301 GGAGTGATCAAAAGCGTATTATGCAACACTTCAGCTCTTAATTTCTGTCATATACAT 360
Db 429 GGAGTGATCAAAAGCGTATTATGCAACACTTCAGCTCTTAATTTCTGTCATATACAT 488
Qy 361 AACTGTATTTCACAGAGATATAAAACTTGAAAATATTCTAATACTAAGCAAGGAATATC 420
Db 489 AACTGTATTTCACAGAGATATAAAACTTGAAAATATTCTAATACTAAGCAAGGAATATC 548
Qy 421 AAGATTTGTGATCTCGGTTTGGCAAAATTTCTGATCCAGAGATGCTTACACCGATTAT 480
Db 549 AAGATTTGTGATCTCGGTTTGGCAAAATTTCTGATCCAGAGATGCTTACACCGATTAT 608
Qy 481 TAGCTACGAGATGCTACCGAGCTCTGAACTTCTTCTGGGAGATACTCAGTAGTGGTTCT 540
Db 609 GTAGCTACGAGATGCTACCGAGCTCTGAACTTCTTCTGGGAGATACTCAGTAGTGGTTCT 668
Qy 541 TCAGTTCGATATATGGGCTATTGGTTGTGTTTTTSCAGAGCTCTCTGACAGGCCCACTG 600
Db 669 TCAGTTCGATATATGGGCTATTGGTTGTGTTTTTSCAGAGCTCTCTGACAGGCCCACTG 728
Qy 601 TGGCTTGAAAATCAGATGAGGACCACTTTTATCTGATATCAGAACACTAG 652
Db 729 TGGCTTGAAAATCAGATGAGGACCACTTTTATCTGATAATCAGAACACTAG 780

RESULT 10

ABX34679

ID ABX34679 standard; cDNA; 1281 BP.

XX

AC ABX34679;

XX

DT 13-FEB-2003 (first entry)

XX

DE Human mddt cDNA SEQ ID 240.

XX

XX MDDT; human; disease detection and treatment molecule polypeptide;
XX anti-inflammatory; immunosuppressive; osteopathic; cycostatic; anti-HIV;
XX haemostatic; nephrotropic; antianemic; antipsoriatic; hepatotropic;
XX gene therapy; protein replacement therapy; cell proliferative disorder;
XX cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;
XX anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
XX Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
XX psoriasis; hepatitis; gene; ss.

XX

CS Homo sapiens.

XX

PN WO200273449-A2.

XX

PC 10-OCT-2002.

XX

XX 27-MAR-2002; 2002WO-US09944.

XX

XX 28-MAR-2001; 2001US-279619P.

XX

XX 29-MAR-2001; 2001US-280067P.

XX

XX 29-MAR-2001; 2001US-280068P.

XX

XX 16-MAY-2001; 2001US-291280P.

XX

XX 17-MAY-2001; 2001US-291829P.

XX

XX 17-MAY-2001; 2001US-291849P.

XX

XX 19-JUN-2001; 2001US-299428P.

XX

XX 20-JUN-2001; 2001US-299776P.

XX

XX 20-JUN-2001; 2001US-300001P.

XX

XX (INCY-) INCYTE GENOMICS INC.

XX

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

XX

XX Duffour CE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SA;

XX

XX Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gersin EH;

XX

XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;

XX

XX Flores V, Marwaha R, Lo A, Lan RY, Urashka NE;

XX

XX WPI; 2003-058431/35.

XX

XX P-PSDB; ABU11689.

XX

XX New purified disease detection and treatment molecule proteins and

XX

XX polynucleotides, useful for diagnosing, treating or preventing cancers

XX

XX (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis

XX

XX or hepatitis -

XX

XX Claim 1; SEQ ID NO 240; 339pp - Sequence Listing; English.

XX

XX This invention describes a novel disease detection and treatment molecule

XX

XX polypeptide (MDPT) which has anti-inflammatory, immunosuppressive,

XX

XX osteopathic, cycostatic, anti-HIV, haemostatic, nephrotropic,

XX

XX antianemic, antipsoriatic and hepatotropic activity. The polynucleotides

XX

XX and the polypeptides of the invention can be used for gene therapy,

XX

XX protein replacement therapy and are useful for treating a variety of

XX

XX diseases or conditions. These polypeptides or polynucleotides are

XX

XX particularly useful for diagnosing, treating or preventing cell

XX

XX proliferative disorders (e.g. cancers including adenocarcinoma,

XX

XX leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's

XX

XX disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's

XX

XX syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or

XX

XX hepatitis. ABX3440-ABX34835 encode the MDDT polypeptides represented in

XX

XX ABU11450-ABU11845, described in the disclosure of the invention.

XX

XX NOTE: The sequence data for this patent did not form part of the printed

XX

XX specification, but was obtained in electronic format from WIPO at

XX

XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1281 BP; 428 A; 191 C; 317 G; 345 T; 0 other;
SQ

Query Katch: 52.2%; Score 543.2; DB 25; Length 1281;
Best Local Similarity 99.5%; Pred. No. 2.8e-135;
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAAAGTATGAAATTTAGCTAAGACTCGAGAGGGTCTTATGGGGTGTATTCAAA 60
DB 704 ATGGAAAGTATGAAATTTAGCTAAGACTCGAGAGGGTCTTATGGGGTGTATTCAAA 763

QY 61 TCAGAAACAAACCTCTGGACAACTAGCTAGCTGTAAAAAATTTGTGGAATTCGAAGAT 120
DB 764 TCGAAGAAACAAACCTCTGGACAACTAGCTAGCTGTAAAAAATTTGTGGAATTCGAAGAT 823

QY 121 GATCCTGTGTGTTAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAACAT 180
DB 824 GATCCTGTGTGTTAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAACAT 883

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGCAAAATGCATTAGTTTTT 240
DB 884 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGCAAAATGCATTAGTTTTT 943

QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300
DB 944 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 1003

QY 301 GGAGTGATCAAAAGGCTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
DB 1004 GGAGTGATCAAAAGGCTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 1063

QY 361 AACTGTATTACAGAGATATAAACTGAAATATTTCTAATACTTAAGCAAGGAATATC 420
DB 1064 AACTGTATTACAGAGATATAAACTGAAATATTTCTAATACTTAAGCAAGGAATATC 1123

QY 421 AAGATTGTGACTTCGGGTTTCACAAATTTCTGATCCAGAGATGCTTACACCGATTAT 480
DB 1124 AAGATTGTGACTTCGGGTTTCACAAATTTCTGATCCAGAGATGCTTACACCGATTAT 1183

QY 481 GTAGCTACGAGATGTTACCGAGCTCTCTGAACTCTTGTGGGAGATACCTCAGTATGTTCT 540
DB 1184 GCAGCTACGAGATGTTACCGAGCTCTCTGAACTCTTGTGGGAGATACCTCAGTATGTTCT 1243

QY 541 TCAGTCTGA 548
DB 1244 TCAGTCTGA 1251

RESULT 11

AAF44669

ID AAF44669 standard; cDNA; 911 BP.

XX

AC AAF44669;

XX

DT 27-MAR-2001 (first entry)

XX

DE Novel protein kinase cDNA, SEQ ID NO: 49.

XX

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
XX immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
XX dermatological; antidiabetic; antifertility; gene therapy; vaccine;
XX immune disorder; cardiovascular disease; neurodegenerative disease;
XX cancer; autoimmune disorder; stroke; inflammatory bowel disease;
XX inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

XX

OS Homo sapiens.

XX

XX WO200073469-A2.

XX

XX 07-DEC-2000.

XX

XX 26-MAY-2000; 2000WO-US14842.

XX

XX

PR 28-MAY-1999; 99US-0136503.
 XX (SUGEN-1) SUGEN INC.
 XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX WPI: 2001-032161/24.
 DR P-PSDB; AAB65642.
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX Example 4; Fig 2; 310pp; English.
 XX The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX Sequence 9:1 BP; 293 A; 182 C; 178 G; 258 T; 0 other:
 SQ Query Match 51.6%; Score 537; DS 22; Length 911;
 Best Local Similarity 87.0%; Pred. No. 1:1e-133;
 Matches 643; Conservative 0; Mismatches 0; Indels 96; Gaps 1;
 285 AAATGGAGTGTGATGGAGTCAAAAGCGTATTATGGCAAACTCAAGCTCTTAA 344
 DB 1 AAATGGAGTGTGATGGAGTCAAAAGCGTATTATGGCAAACTCAAGCTCTTAA 60
 345 TTCTGTCTATACATGATGATTTTCACAGAGATATAAAGCTGAAATATCTAATAC 404
 DB 61 TTCTGTCTATACATGATGATTTTCACAGAGATATAAAGCTGAAATATCTAATAC 120
 405 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTCGACAAATTTGATTCAGGAGA 464
 DB 121 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTCGACAAATTTGATTCAGGAGA 180
 465 TGGCTACACCGAATATGTAGCTACAGATGTTACCGAGCTCTGAACTTTTGTGGGAGA 524
 DB 181 TGGCTACACCGAATATGTAGCTACAGATGTTACCGAGCTCTGAACTTTTGTGGGAGA 240
 525 TACTCAGTATGTTCTTCAGTCGATATGGCTATTGTTGTGTTTTCAGAGCTCT 584
 DB 241 TACTCAGTATGTTCTTCAGTCGATATGGCTATTGTTGTGTTTTCAGAGCTCT 300
 585 GACAGCCAGCCACTGTGGCTTGGAAATTCAGATGTGACCAACTTTATCTGATAATCAG 644
 DB 301 GACAGCCAGCCACTGTGGCTTGGAAATTCAGATGTGACCAACTTTATCTGATAATCAG 360
 645 AACACTAGTAGAGACGGGGTTTCGCCATCTTCACAGGCTGGTCTCGAACTCTTGACGTC 704
 DB 361 AACACT-----
 705 AAGTGATCCACCTGCGGTAGCTCTCAAGTGTCTGAAATTCAGGAAATTAATCCCAAG 764
 DB 367 -----AGGAAATTAATCCCAAG 384
 765 ACATCAATCAATCTTTAAAGTAAGCGGTTTTTCATGGCATCTAGTATCTGAGCGAGA 824
 QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 385 ACATCAATCAATCTTTAAAGTAACGGGTTTTCCATGGCATCAGTATACCTGAGCAGA 444
 QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 445 AGACATGGAAACTCTTGAAGGAAAGTTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCAT 504
 QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 505 GAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACTGTTTCCCACTCTCTGGAGAG 564
 QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 565 CTCTACTTTGATTTCTTCAAGAGGCCCAAAATTTAAAGGAAAGACGTTAATGAAGGAAG 1004
 QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1005 AAACAGAGAGCGCCACACAG 1023
 DB 625 AAACAGAGAGCGCCACACAG 643
 AC AAD3812;
 DT 19-JUN-2001 (first entry)
 XX Human Kinase cDNA #1.
 XX Human; kinase; gene therapy; bioreactor; mental disorder;
 XX Biological disorder; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT 1.561
 FT /tag= a
 FT /product= "Human kinase #1"
 FT /note= "The coding region does not include stop codon"
 XX WC200123579-A1.
 XX 05-APR-2001.
 XX 27-SEP-2000; 2000WO-US26621.
 XX 28-SEP-1999; 99US-0156511.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Doncho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX WPI: 2001-266166/27.
 XX P-PSDB; AAB00490.
 XX New isolated human kinase polynucleotide useful for generating
 PT antibodies, as reagents in diagnostic assays and for screening for
 PT compounds useful for treating mental, biological or medical diseases -
 XX Disclosure; Page 27; 38pp; English.
 XX The present sequence is a cDNA encoding novel human protein
 CC (NHP) known as human kinase. The human kinases share structural
 CC similarity with animal kinases, more particularly serine or
 CC threonine protein kinases. Human kinase cDNA is useful for the
 CC detection of mutant human kinase for the diagnosis of disease,
 CC and also as a therapeutic. It is useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic
 CC manifestations perturbing the normal function of NHP in the
 CC body. The NHP nucleotide sequences are useful for generation of
 CC antibodies, as reagents in diagnostic assays, for the
 CC identification of other cellular gene products related to human

Db 301 GGAGTGTATCAAAAGCGTATTATGCAAGCACTTCAAGCTCTTAAT-TCTGTCAATATGAT 360
QY 361 AACTGTATTACACAGAGTATTAACACCTGAAATATCTTAATACCTAAGCAAGCAATATC 420
Db 361 AACTGTATTACACAGAGTATTAACACCTGAAATATCTTAATACCTAAGCAAGCAATATC 420
QY 421 AAGATTCTGACTTCGGGTTTGCACAAATCTGATT 456
Db 421 AAGATTCTGACTTCGGGTTTGCACAAATCTGAGT 456
RESULT 14
AAF44670
ID AAF44670 standard; cDNA; 2615 BP.
XX
AC AAF44670;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase cDNA, SEQ ID NO: 50.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiac; renal; antiinflammatory; antilethemic;
KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ps.
XX
OS Mus musculus.
XX
FN WO2000073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; ZOO-WO-US14842.
XX
PR 28-MAY-1999; 99US-0136503.
XX
FA (SUGEN-) SUGEN INC.
XX
P- Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
WPI: 2001-032:6:/04.
DR P-PSDB: AAB65643.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers.
XX
PS Disclosure; Fig 2; 310pp; English.
XX
CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 2615 BP; 742 A; 580 C; 585 G; 708 T; 0 other;

Query Match 43.5%; Score 453.2; DB 22; Length 2615;

Best Local Similarity 73.1%; Pred. No. 4.6e-111;
Yatches 658; Conservative 0; Mismatches 143; Indels 99; Gaps 2;
QY 135 GAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAGACATCCAAATCTTGTGAA 194
Db 1 GAAATAGCCCTGGCGGAAATCCGTATGCTGAAG--TTGAAACACCAACCTCGTGAA 57
QY 195 CCTCATCGAGGTGTTCCAGGAGAAAGAAATGCATTCTAGTTTGTGAATACTGTGATCA 254
Db 58 CCTCATCGAGGTGTTCAAGAGAAAGAAAGATGATCTAGTTTTTGTAGTACTGTGATCA 117
QY 255 TACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGATCGAGTGATCAAAAG 314
Db 118 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGTGAGTGAATTAAG 177
QY 315 CGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACATACTGTATTCCAG 374
Db 178 TGCTGTATGGCAACCTTCAAGCCCTTAACTTCTGTCAAGCACAATTTGTATTATCG 237
QY 375 AGATATAAACCCTGAAATATTTCTAATACTAAGCAAGGAATAATCAAGATTTGTGACTT 434
Db 238 GGATGTAACCTGAAACATCTTAATAACCAAGCAGGATGATAAAGATTTGTGACTT 297
QY 435 CGGTTTGCACAAATCTGTATTCAGGAGATGCTACACCGATTTATGTAGTACGAGATG 494
Db 298 TCGATTTTGCACGAATCTTAATTTCCAGGAGAGCGCTACACAGACTATGTTGCCACCGTG 357
QY 495 GTACCGAGCTCTCAACTCTTTGTGGGAGATCTCAGTATGTTCTTCAGTCGATATATG 554
Db 358 GTACCGAGCCCGAATCTTCTGTGGAGACACGAAGTACGGTCTCTGTAGACGTGTG 417
QY 555 GGTATTGTTGTTCTTTTTCAGAGCTCTTGACAGGCCAGCCACTGTGGCTGGAAATC 614
Db 418 GGCCTCTGGCTGTGTTTTCAGAGACTCTTGAGGGTCAGCCACTCTGGCCGGGAAATC 477
QY 615 AGATGTGGACCAACTTTATCTGATTAATCAGAACACTAGTAGAGACGGGTTTCCCATGT 674
Db 478 CGAGTGGACCAAGCTTTTACCTGATCATCAGGACGTTG----- 514
QY 675 TGACCAGGCTGCTCGAACTCTTGAGCTCAAATGATCCACCTGCCCTAGCCTCTCAAG 734
Db 515 ----- 514
QY 735 TGCTGGAATTACAGAGAAATTAATCCCAAGACATCAATCAATCTTTAAAAGTAACGGTT 794
Db 515 -----GGGAAGCTGATTCCCAAGACCACTCTATCTTTAGGAGTAACCAAGTT 561
QY 795 TTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTCTTGAGGAAAGTTTCTC 854
Db 562 TTTCCGGCGCATCAGCATACCTGAACCAAGAGGACATGAGACTCTTGAGAGAAATTTCTC 621
QY 855 AGATGTTTCATCTGTGCTGTGAACCTTCATGAAGGGGTGCTGAAGATGAATCCAGATGA 914
Db 622 AAATGTTTCAGCTGTGGCTTTAAGTTTTCATGAAGGGATGCTGAAGATGAATCCTGATGA 681
QY 915 CAGATTAACTGTTTCCCAACTCTTGAGAGCTCTTACTTGTATCTTTTCAAGAGGCCCA 974
Db 682 GAGGCTGACCTGTGCCCAGCTGCTGGACAGTGCCTACTTTGAGTCTTTTCAAGAGGATCA 741
QY 975 AATTAAGAGAAAGCAAGCTTAATGAAGCAAGAAACAGAGACGCCAACAGGTACTTCCGCT 1034
Db 742 AATGAAGAGAAAGCCCGCAGTGAGGGGAGAGCCGAAGGCGCCAGCAGATCAACTGCT 801
RESULT 15
ABZ77165
ID ABZ77165 standard; cDNA; 1086 BP.
XX
AC ABZ77165;
XX
DT 07-MAY-2003 (first entry)
XX
DE Human protein kinase encoding cDNA SEQ ID NO:79.

DB	463	ACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACTCAG	522
QY	532	TATGGTCTTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTCGACAGTCCCTGCACAGCG	591
DB	523	TATGGTCTTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTCGACAGTCCCTGCACAGCG	582
QY	592	CAGCACGTGGCTCGAAATCAGATGTGGACAACTTTATCTGATAATCAGAACACTA	651
DB	583	CAGCACGTGGCTCGAAATCAGATGTGGACAACTTTATCTGATAATCAGAACACT	641
QY	652	GTAGAGACGGGGTTTCGCCATGTTGACACGGCTGGTCTCGAACTCTTGAGCTCAAGTGAT	711
DB	642	-----	641
QY	712	CGACCTGCGCTAGCCTCTCAAGTGTCTGGAATTACAGGAAAAATTAATCCACAGACATCAA	771
DB	642	-----AGGAAAAATTAATCCACAGACATCAA	666
QY	772	TCAATCTTTAAAGTAACGGGTTTTCCATGGCATCAGTATACCTCGCCGACAGAACATG	831
DB	667	TCAATCTTTAAAGTAACGGGTTTTCCATGGCATCAGTATACCTCGCCGACAGAACATG	726
QY	832	GAACCTCTTGAGGAAAAAGTTCTCAGATGTTTCATCCCTGTGCTCTGAACCTCATGAAG	888
DB	727	GAACCTCTTGAGGAAAAAGTTCTCAGATGTTTCATCCCTGTGCTCTGAACCTCATGAAG	783

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:42:01 / Search time 4078.56 Seconds
(without alignments)
10441.655 Million cell updates/sec

Title: US-09-671-050-9

Perfect score: 1041

Sequence: 1 atggaaaagatgaaaaatt.....aggtaacttcgctcaaaagt 1041

Scoring table: IDENTITY NUC

Gapop 10_0, Gapext 1.0

Searched: 2888711 seqs, 2045481336 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl:

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.ot.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.ham.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1041	100.0	1041	6	AX107720	Sequence
2	1025	98.5	1068	6	AX107714	Sequence
3	839	80.6	945	6	AX107722	Sequence
4	823	79.1	972	6	AX107716	Sequence
5	762.8	72.3	1678	6	AX698820	Sequence
6	757	72.7	1083	6	AX166534	Sequence
7	709	68.1	1819	6	AX107724	Sequence
8	645.6	62.0	687	6	AX286069	Sequence
9	645.6	62.0	882	6	AX286067	Sequence
10	537	51.6	911	6	AX056404	Sequence
11	454.4	43.7	561	6	AX107712	Sequence
12	454.4	43.7	594	6	AX107718	Sequence
13	453.2	43.5	2615	6	AX056405	Sequence
14	343.8	33.0	1363	9	HSTHPKB	X66358 H. sapiens m
15	343.8	33.0	1612	6	AX399975	Sequence
16	342.8	32.9	2952	3	AX112442	Sequence
17	251.8	24.2	1701	6	AR221995	Sequence
18	251.8	24.2	3080	6	AR221994	Sequence
19	251.8	24.2	3153	4	AB029045	Oryctolag
20	243.8	23.4	1518	10	AB029067	Mus muscu
21	243.8	23.4	3988	10	AB029066	Mus muscu
22	243.8	23.4	4022	10	AB029065	Mus muscu
23	241.2	23.2	1993	9	HSU35146	U35146 Homo sapien
24	202.6	19.5	1513	9	AF130372	Homo sapi
25	202.6	19.5	1733	6	AX166533	Sequence
26	202.6	19.5	2146	9	BC041799	Homo sapi
27	201	19.3	1974	3	AF326965	Trypanoso
28	200.6	19.3	1620	10	BC028871	Mus muscu
29	199	19.1	80362	9	AC079615	Homo sapi
30	191	18.3	1732	10	AF112183	Rattus no
31	191	18.3	1738	10	AF112184	Rattus no
32	186.6	17.9	101340	2	AC018104	Drosophil
33	186.6	17.9	190668	3	AC007419	Drosophil
34	186.6	17.9	260731	3	AE003611	Drosophil
35	177.2	17.0	3399	9	AY217744	Homo sapi
36	177.2	17.0	3399	9	HSTXP3	Homo sapien
37	177.2	17.0	3441	9	BC036091	Homo sapi
38	175.2	16.8	1866	6	AX056407	Sequence
39	166	15.9	960	3	TCU069958	Trypanosoma
40	163.8	15.7	903	8	AF026545	Pneumocys
41	162	15.6	1476	6	AR087351	Sequence
42	162	15.6	1476	6	AR281916	Sequence
43	162	15.6	1476	9	HSCDK2MR	X61622 H. sapiens C
44	162	15.6	1635	6	A84497	Sequence 11
45	162	15.6	1635	6	A84499	Sequence 11

ALIGNMENTS

RESULT 1
AX107720
LOCUS AX107720 1041 bp linear PAT 30-APR-2001
DEFINITION Sequence 9 from Patent WO0123579.
ACCESSION AX107720
VERSION AX107720.1 GI:13923201
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Donoho G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and

Sande, A.T.

Human Kinase proteins and polynucleotides encoding the same


```
Db 421 AAGATTGTGACTTCGGGTTGCACAAATTCGATTCCAGAGATGCCTACACCGATTAT 480
Qy 481 GTAGCTACGAGATGGTACCGAGCTCTGAACCTCTGTGGGAGATACCTCAGTATGGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCTGAACCTCTGTGGGAGATACCTCAGTATGGTTCT 540
Qy 541 TCAGTCCGATATATGGGCTATTGGTTGTGTTTTCGAGAGCTCTGACAGGCGCAGCCACTG 600
Db 541 TCAGTCCGATATATGGGCTATTGGTTGTGTTTTCGAGAGCTCTGACAGGCGCAGCCACTG 600
Qy 601 TGCCCTGGGAAATCAGATGTGGACCACTTATCTGATNATCAGACACTAGTAGAGAG 660
Db 601 TGCCCTGGGAAATCAGATGTGGACCACTTATCTGATNATCAGACACTAGTAGAGAG 660
Qy 661 GGGTTTCGCCCATGTTGACCAAGGCTGTCTCGAACTCTTGACGTCGAGTGCACCTGCC 720
Db 661 GGGTTTCGCCCATGTTGACCAAGGCTGTCTCGAACTCTTGACGTCGAGTGCACCTGCC 720
Qy 721 GTAGCCTCTCAAGTGTGGAAATCAGGAAATTAATCCGAGACATCAATCAATCTTT 780
Db 721 GTAGCCTCTCAAGTGTGGAAATCAGGAAATTAATCCGAGACATCAATCAATCTTT 780
Qy 781 AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCGAGAGACATGAAACTCTT 840
Db 781 AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCGAGAGACATGAAACTCTT 840
Qy 841 GAGGAAAGTTCTCAGATGTTCATCTGCTGGCTCGAATTCATGAAGGGGTCTGTAAG 900
Db 841 GAGGAAAGTTCTCAGATGTTCATCTGCTGGCTCGAATTCATGAAGGGGTCTGTAAG 900
Qy 901 ATGAATCCAGATCAGAGATTAACCTGCTCCCACTCTGAGAGCTCTGATTTCT 960
Db 901 ATGAATCCAGATCAGAGATTAACCTGCTCCCACTCTGAGAGCTCTGATTTCT 960
Qy 961 TTTCAAGAGCCCCAAATTAAGAAAGACACGTATGAAGGAGAAAGACAGAGAGCGCAA 1020
Db 961 TTTCAAGAGCCCCAAATTAAGAAAGACACGTATGAAGGAGAAAGACAGAGAGCGCAA 1020
Qy 1021 CASGT 1025
Db 1021 CAGGT 1025

RESULT 3
AX107722
LOCUS AX107722 945 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 11 from Patent WO0123579.
ACCESSION AX107722
VERSION AX107722.1 GI:13923202
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 11 05-APR-2001;
FEATURES
Location/Qualifiers
source 1..945
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 165 c 202 g 255 t
ORIGIN
Query Match 80.6%; Score 839; DB 6; Length 945;
Best Local Similarity 90.8%; Pred. No. 3.3e-193;
Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;
Qy 1 ATGGAAGATGAAATTTAGCTAGAGAGGCTCTTATGGGGTTGTATTCAA 60
```

```
1 ATGGAAGATGAAATTTAGCTAGAGAGGCTCTTATGGGGTTGTATTCAA 60
61 TGCGAAAAAACAACCTCTCGCAACAAGTAGCTGTATAAAAAATTTGTGGAATCTCAAGAT 120
61 TGCGAAAAAACAACCTCTCGCAACAAGTAGCTGTATAAAAAATTTGTGGAATCTCAAGAT 120
121 GATCCTGTTGTTAAGAAAAATACGACTAGAGAAATACGTTATGTTCAAGCAATTTAAACAT 180
121 GATCCTGTTGTTAAGAAAAATACGACTAGAGAAATACGTTATGTTCAAGCAATTTAAACAT 180
181 CCAATCTTTGTGAACCTCATCGAGGTGTTACGAGAAAAGGAAAAATGCATTTAGTTTTT 240
181 CCAATCTTTGTGAACCTCATCGAGGTGTTACGAGAAAAGGAAAAATGCATTTAGTTTTT 240
241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
301 GSAGTGAATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
301 GSAGTGAATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
361 AACTGTATTACAGAGATATAAACTGAAATATTTCTAATACTAAGCAGGAGTAATC 420
361 AACTGTATTACAGAGATATAAACTGAAATATTTCTAATACTAAGCAGGAGTAATC 420
421 AAGATTGTGACTTCGGGTTTGCACAAAATTTGATTCAGGAGATGCTTACACCCATTAT 480
421 AAGATTGTGACTTCGGGTTTGCACAAAATTTGATTCAGGAGATGCTTACACCCATTAT 480
481 GTAGCTACGAGATGGTACCGAGCTCTGAACTCTTGTGGGAGATACCTCAGTATGGTTCT 540
481 GTAGCTACGAGATGGTACCGAGCTCTGAACTCTTGTGGGAGATACCTCAGTATGGTTCT 540
541 TCAGTCGATATATGGCTATTCGGTTGTGTTTTGTCAGAGCTCTTACAGGCGCAGCCACTG 600
541 TCAGTCGATATATGGCTATTCGGTTGTGTTTTGTCAGAGCTCTTACAGGCGCAGCCACTG 600
601 TGCCCTGGGAAATCAGATGTGGACCACTTTATCTGATNATCAGAACTAGTAGAGAG 660
601 TGCCCTGGGAAATCAGATGTGGACCACTTTATCTGATNATCAGAACTAGTAGAGAG 660
661 GGGTTTCGCCCATGTTGACCAAGGCTGTCTCGAACTCTTGACGTCGAGTGCACCTGCC 720
661 GGGTTTCGCCCATGTTGACCAAGGCTGTCTCGAACTCTTGACGTCGAGTGCACCTGCC 720
651 -----
721 GTAGCCTCTCAAAGTGTCTGGAATTAACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
781 AAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCGAGAGACATGGAACCTTT 840
685 AAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCGAGAGACATGGAACCTTT 744
841 GAGGAAAGTTCTCAGATGTTCATCTGCTGGCTCTGAACTCTCATGAAGGGGTGCTGAAG 900
745 GAGGAAAGTTCTCAGATGTTCATCTGCTGGCTCTGAACTCTCATGAAGGGGTGCTGAAG 804
901 ATGAATCCAGATCAGAGATTAACCTGTTCCCACTCTGAGAGCTCTTACTTTGATTTCT 960
805 ATGAATCCAGATCAGAGATTAACCTGTTCCCACTCTGAGAGCTCTTACTTTGATTTCT 864
961 TTTCAAGAGCCCCAAATTAAGAAAGACACGTATGAAGGAGAAAGACAGAGAGCGCAA 1020
865 TTTCAAGAGCCCCAAATTAAGAAAGACACGTATGAAGGAGAAAGACAGAGAGCGCAA 924
1021 CAGGTACTTCCGCTCAAAAGT 1041
925 CAGGTACTTCCGCTCAAAAGT 945
```

RESULT 4

```
AX107716
LOCUS AX107716 972 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 5 from Patent WO0123579.
ACCESSION AX107716
VERSION AX107716.1 GI:13923199
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 5 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..972
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 172 c 214 g 263 t
ORIGIN
Query Match 79.1%; Score 823; DB 6; Length 972;
Best Local Similarity 90.6%; Pred. No. 2.5e-189;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;
QY 1 ATGGAAAGTATGAAATATGCTAGAGCTGGAGAGGGCTTATGGGTGTATTCAA 60
DB 1 ATGGAAAGTATGAAATATGCTAGAGCTGGAGAGGGCTTATGGGTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTGTATAAAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTGTATAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCTGTGTTAAGAAATATGACATAGAGAAATACGTATGTTGAAGCAATTAACAT 180
DB 121 GATCTGTGTTAAGAAATATGACATAGAGAAATACGTATGTTGAAGCAATTAACAT 180
QY 181 CCAAATCTGTGAACTCATCGAGTGTTCAGGAGAAAGAAATGCAATTAAGTTT 240
DB 181 CCAAATCTGTGAACTCATCGAGTGTTCAGGAGAAAGAAATGCAATTAAGTTT 240
QY 241 GAATCTGTGATCATACATTTTAATGAGCTGGAGAAACCCAAATCGAGTTGCTGAT 300
DB 241 GAATCTGTGATCATACATTTTAATGAGCTGGAGAAACCCAAATCGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATATGGCAACACCTTCAAGCTCTAAATTTCTGTCAATACAT 360
DB 301 GGAGTGATCAAAAGCGTATATGGCAACACCTTCAAGCTCTAAATTTCTGTCAATACAT 360
QY 361 AACTGTATTTCAGAGATATAAAACCTGAAATATTCATAAATAAGCAAGGAATATC 420
DB 361 AACTGTATTTCAGAGATATAAAACCTGAAATATTCATAAATAAGCAAGGAATATC 420
QY 421 AGATTGTGACTTCGGGTTCGCAAAATTCGATTCAGGAGATGCTACACCGATAT 480
DB 421 AGATTGTGACTTCGGGTTCGCAAAATTCGATTCAGGAGATGCTACACCGATAT 480
QY 481 GTAGCTACAGATGTTACGAGCTCTGAACTCTTGTGGAGATACCTCAGTATGGTCT 540
DB 481 GTAGCTACAGATGTTACGAGCTCTGAACTCTTGTGGAGATACCTCAGTATGGTCT 540
QY 541 TCAGTCGATATATGGCTATATGGTGTGTTTTCAGAGCTCTTCAGAGCCAGCCACTG 600
DB 541 TCAGTCGATATATGGCTATATGGTGTGTTTTCAGAGCTCTTCAGAGCCAGCCACTG 600
QY 601 TGGCTTGGAAATCAGATGTGGACCAACTTTATCTGTATAATCAGAACTAGTAGAGCG 660
DB 601 TGGCTTGGAAATCAGATGTGGACCAACTTTATCTGTATAATCAGAACTAGTAGAGCG 660
QY 661 GGTTTCGCCATGTTGACCAAGCTGGTCTCGAACTCTTTGACGTCAAGTGATCCACCTGCC 720
DB 661 GGTTTCGCCATGTTGACCAAGCTGGTCTCGAACTCTTTGACGTCAAGTGATCCACCTGCC 720
DB 651 ----- 650
QY 721 GTAGCTCTCAAGTGTCTGGAATTAATCCCAAGACATCAATCAATCTTT 780
DB 721 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 781 AAAAGTAAAGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTCT 840
DB 781 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 841 GAGGAAAGTTCTCAGATGTTTCATCTCTGCTCTGAATCTCATGAAGGGGTCTCTGAAG 900
DB 841 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 901 ATGAATCCAGATGACAGATTAACTCTGTTCCCACTCTCTGAGAGCTCTCTATTGATTCT 960
DB 901 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 961 TTTCAAGAGGCCCAATTTAAAGAAAGAACAGCTAATGAAGAAAGAAACAGAACGCCAA 1020
DB 961 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 1021 CAGGT 1025
DB 925 CAGGT 929
RESULT 5
AX698820 1678 bp DNA linear PAT 02-APR-2003
LOCUS AX698820
DEFINITION Sequence 6 from Patent WO2079473.
ACCESSION AX698820
VERSION AX698820.1 GI:29499608
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Panzer,S.R., Lirio,S.E., Altus,C.M., Dufour,G.E., Hillman,J.L., Jones,A.L., Dam,T.C., Liu,T.F., Harris,B., Flores,V., Daffo,A., Marwaha,R., Cher,A.J., Chang,S.C., Gerstin,J.E., Peralta,C.H., David,N.H. and Lewis,S.A.
TITLE Molecules for diagnostics and therapeutics
JOURNAL Patent: WO 02079473-A 6 10-OCT-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1678
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LI:058298.1:2001JUN12"
BASE COUNT 541 a 308 c 336 g 493 t
ORIGIN
Query Match 73.3%; Score 762.8; DB 6; Length 1678;
Best Local Similarity 90.3%; Pred. No. 1e-174;
Matches 925; Conservative 0; Mismatches 2; Indels 101; Gaps 6;
QY 1 ATGGAAGATGATGAAATTAAGCTAGAGCTGGAGAGGGTCTTATGGGTGTATTCAA 60
DB 131 ATGGAGAAGTATGAAATTAAGCTAAGACTGGAGAGGGTCTTATGGGTGTATTCAA 190
QY 61 TGCAGA-AACAAACCTCTGCAGCAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAGA 199
DB 61 TGCAGATAAACAAACCTCTGCAGCAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAGA 250
QY 120 TGATCTCTGTTTGAAGAAATAGCACCTAAGAGAAATACGTA-TGTTGAAGCAATTAAC 178
DB 251 TGATCTCTGTTTGAAGAAATAGCACCTAAGAGAAATACGTA-TGTTGAAGCAATTAAC 310
QY 179 ATCCAATCTTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 238
DB 179 ATCCAATCTTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 238
```

311 ATCCAAATCTTGTAACCTCATCGAGGTGTTTCAGAGAGAAAGGAAATGCAATTTAGTTT 370
312 TTGAATACCTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTG 298
313 TTGAATACCTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTG 430
314 ATGGAGTATCAAAAGCGTATTATCGGCAACACTTCAAGCTCTTAAATTTCTGTCATATAC 358
315 ATGGAGTATCAAAAGCGTATTATCGGCAACACTTCAAGCTCTTAAATTTCTGTCATATAC 490
316 ATAACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATAACAAAGCAAGATATA 438
317 ATAACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATAACAAAGCAAGATATA 550
318 TCAAGATTGTGACCTTCCGGTTTGCAAAATCTTGATCTCCAGGAGATGCGCTACACCGATT 478
319 TCAAGATTGTGACCTTCCGGTTTGCAAAATCTTGATCTCCAGGAGATGCGCTACACCGATT 610
320 TCAAGATTGTGACCTTCCGGTTTGCAAAATCTTGATCTCCAGGAGATGCGCTACACCGATT 535
321 ATGTAGCTA-CGAGATGGTACCGAGCT-CCTGAACTTCTGCGAGATACT-CAGTATG 670
322 ATGTAGCTA-CGAGATGGTACCGAGCT-CCTGAACTTCTGCGAGATACT-CAGTATG 595
323 GTCTCTGAGTATATGGGCTATGGTCTGTTTGGAGACTCTGACAGGCCAGC 730
324 GTCTCTGAGTATATGGGCTATGGTCTGTTTGGAGACTCTGACAGGCCAGC 655
325 CACTGTGGCTGGAATACAGATGTGACCAACTTTATCTGATATCAGAACACTAGTAG 785
326 CACTGTGGCTGGAATACAGATGTGACCAACTTTATCTGATATCAGAACACTAGTAG 715
327 CACTGTGGCTGGAATACAGATGTGACCAACTTTATCTGATATCAGAACACTAGTAG 785
328 AGACGGGTTCGCGCATGTGACCAAGGTGCTCTGAACTCTTGACGTCAAGTCAATCCAC 775
329 CTGCGGTAGCTCTCAAGTGTGGAATTACAGGAAATTAATCCCAAGACATCAATCAA 834
330 CTGCGGTAGCTCTCAAGTGTGGAATTACAGGAAATTAATCCCAAGACATCAATCAA 874
331 TCTTTAAAGTAACGGTTTCCATGTCATCAGTATACCTGAGCCAGACATGGA 895
332 TCTTTAAAGTAACGGTTTCCATGTCATCAGTATACCTGAGCCAGACATGGA 934
333 CTCTTGAAGAAAGTCTCAGATGTTCACTCTGCGCTCTGAACTTCTAAGGGGTGTC 955
334 TGAAGATGAATCCAGATGACAGATTAACTGTCTCCAACTCTCTGAGAGCTCTTACTTTG 994
335 TGAAGATGAATCCAGATGACAGATTAACTGTCTCCAACTCTCTGAGAGCTCTTACTTTG 1015
336 ATCTTTTCAAGAGGCCAAATTAAGAAAGACGTTAAGAGGAAAGAAACAGAGAC 1054
337 ATCTTTTCAAGAGGCCAAATTAAGAAAGACGTTAAGAGGAAAGAAACAGAGAC 1023
338 GCACACAG 1062
339 GCACACAG 1062

RESULT 6
AX166534
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX166534
Sequence 25 from Patent WO0138503.
AX166534
AX166534.1 GI:14546879
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Plozman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,

Flanagan,P. and Clary,D.S.
Novel human protein kinases and protein kinase-like enzymes
Patent: WO 0138503-A 25 31-MAY-2001;
Sugen, Inc. (US)
Location/Qualifiers
1..1083
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 366 a 199 c 225 g 293 t
ORIGIN
Query Match 72.7%; Score 757; DB 6; Length 1093;
Best Local Similarity 86.1%; Pred. No. 2.7e-173;
Matches 927; Conservative 0; Mismatches 0; Indels 150; Gaps 2;
QY 1 ATGGAAGATGTAAGAAATTTAGCTAAGACTGGAGAGGCTCTTATGGGTTGTATTCAAA 60
DB 1 ATGGAAGATGTAAGAAATTTAGCTAAGACTGGAGAGGCTCTTATGGGTTGTATTCAAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAAGTAGTACTCTTAAAAAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAAAGTAGTACTCTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTTGAAGAAATAGCACTAAGAGAAATAGTATGTTGAAGCAATTAAGACAT 180
DB 121 GATCCTGTTTGAAGAAATAGCACTAAGAGAAATAGTATGTTGAAGCAATTAAGACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGCTGTTTCAAGAGAAAGGAAATGCAATTTAGTTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGCTGTTTCAAGAGAAAGGAAATGCAATTTAGTTTT 240
QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTATCAAAAGGCTATTATGGCAACACTTCAAGCTCTTAAATTTCTGTCATATACAT 360
DB 301 GGAGTATCAAAAGGCTATTATGGCAACACTTCAAGCTCTTAAATTTCTGTCATATACAT 360
QY 361 AACTCTATTTCAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
DB 361 AACTCTATTTCAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
QY 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGA----- 454
DB 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGAAGTAACTTCAAGTGGAGAAATGAC 480
QY 455 -----TTCAGGAGATGCTACACCGATTATGTAGCT 486
DB 481 TTACCTGATCTGATTGTCATTTTCAGTTCAGAGATGCTTACACCGATTATGTAGCT 540
QY 487 ACGAGATGGTACCGAGCTCTGAACTTTCTGAGGAGATCTCAGTATGTTCTTCAGTC 546
DB 541 ACGAGATGGTACCGAGCTCTGAACTTTCTGAGGAGATCTCAGTATGTTCTTCAGTC 600
QY 547 GATATATGGGCTATTGGTTGTTTTCAGAGCTCTGACAGCCAGCCACTGTGGCCT 606
DB 601 GATATATGGGCTATTGGTTGTTTTCAGAGCTCTGACAGCCAGCCACTGTGGCCT 660
QY 607 GAAAAATCAGATGGGCAACTTTTATCTGATATCAGAACACTAGTAGAGCGGGTTT 666
DB 661 GAAAAATCAGATGGGCAACTTTTATCTGATATCAGAACACT----- 704
QY 667 CCGCATGTTGACCGGCTGCTCGAAGCTTTGAGCTCAAGTGAATCCACCTGCGGTAGCC 726
DB 705 ----- 704
QY 727 TCTCAAGTGTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTTAAAGT 786
DB 705 -----AGGAAATTAATCCCAAGACATCAATCAATCTTTAAAGT 744
QY 787 AAGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTCTTTGAGGAA 846

```
Db 745 AACGGGTTTTCATGATCAGTATACCTGAGCCAGAACATGAATCTGTGGAA 864
Qy 847 AAGTTCCTCAGATGTTTCCTGTGGCTCTGAACCTTCAATGAAGGGTGTCTGAAGATGAAT 906
Db 805 AAGTTCCTCAGATGTTTCCTGTGGCTCTGAACCTTCAATGAAGGGTGTCTGAAGATGAAT 864
Qy 907 CCAGATGACAGATTAACTGTTCCCACTCTCTGAGAGCTTCTACTTTGATTTCTTTCA 966
Db 865 CCAGATGACAGATTAACTGTTCCCACTCTCTGAGAGCTTCTACTTTGATTTCTTTCA 924
Qy 967 GAGGCCCAAAATTAAGAGAAAGCAGCTTAATGAAGAGAAAGAAACAGAGCCCAACAG 1023
Db 925 GAGGCCCAAAATTAAGAGAAAGCAGCTTAATGAAGAGAAAGAAACAGAGCCCAACAG 981

RESULT 7
AX107724
LOCUS AX107724 1819 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 13 from Patent WO0123579.
ACCESSION AX107724
VERSION AX107724.1 GI:13923203
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 13 05-APR-2001;
LEXICON Lexicon Genetics Incorporated [US];
FEATURES
    Location/Qualifiers
        1..1819
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT 539 a 368 c 406 g 505 t 1 others
ORIGIN

Query Match 68.1%; Score 709; DB 6; Length 1819;
Best Local Similarity 77.0%; Pred. No. 1.2e-161;
Matches 1025; Conservative 0; Mismatches 0; Indels 306; Gaps 1;

Qy 1 ATGGAAAAGTATGAAAATTAAGCTTAAGACTGGAGAGGGCTTATGGGTGTATTCAAA 63
Db 138 ATGGAAAAGTATGAAAATTAAGCTTAAGACTGGAGAGGGCTTATGGGTGTATTCAAA 197
Qy 61 TGCAGAAACAAACCTCTGGACAGTACTAGCTGTAAAGAAATTTGGGAAATCTGAAGAT 120
Db 198 TGCAGAAACAAACCTCTGGACAGTACTAGCTGTAAAGAAATTTGGGAAATCTGAAGAT 257
Qy 121 GATCTCTGTGTTTAAGAAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAACAT 180
Db 258 GATCTCTGTGTTTAAGAAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAACAT 317
Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAGAAAGCAATGCAATAGTTT 240
Db 318 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAGAAAGCAATGCAATAGTTT 377
Qy 241 GAATACCTGTGATCATACATTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT 300
Db 378 GAATACCTGTGATCATACATTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT 437
Qy 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTTTAATTTCTGTCTATACAT 360
Db 438 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTTTAATTTCTGTCTATACAT 497
Qy 361 AACTGTATTTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
Db 498 AACTGTATTTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 557
```

```
Qy 421 AGATTTGTGACTTCGGTTTSCAAAAATTCG----- 453
Db 559 AGATTTGTGACTTCGGTTTSCAAAAATTCGAGTTGGACTTCATCTTTCTCTGTGTGCC 617
Qy 454 ----- 453
Db 518 TCCTTGATTGGCTTAATAGTTGACCTTCTGAAATTC-----TTTCTGCCAAATTCAGAGA-----TTT 677
Qy 454 ----- 453
Db 678 CTCTGCTTTGGATTCATTTCTGCACAGGTGTTTCACATGGGGGCCAGGCTCATCTCGA 737
Qy 454 ----- 453
Db 738 ACTTCTGGCCTCAAGTGATCTCTCCACCTCGGCTCCCAAGTGTGGATTSCAAGTGTG 797
Qy 454 ----- 453
Db 796 AGCCACCGTCCCGAGCCAGATTTTCAAAACAATACTACTGAGAGCTCAAAAGATTGTTT 857
Qy 454 -----ATTCCAGGAGATGCCTACACC 474
Db 958 TTAGTGGGACACAAATTCGAACAAATTCCTGGAAGCGCATTCAGGAGATGCCTACACC 917
Qy 475 GATTATGCTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATACTCAGTAT 534
Db 918 GATTATGCTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATACTCAGTAT 977
Qy 535 GGTCTCTCAGTTCGATATGCGCTATTGGTATTGTTGTTTGTGCAGAGCTCTCTACAGCCAG 594
Db 978 GGTCTCTCAGTTCGATATGCGCTATTGGTATTGTTGTTTGTGCAGAGCTCTCTACAGCCAG 1037
Qy 595 CCACGTGTGGCTGGAATAACAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA 654
Db 1038 CCACGTGTGGCTGGAATAACAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA 1097
Qy 555 GAGAGGGGGTTTCGCCATGTTGACCGAGGTGGTCTCGAACTCTTGACGTCAAGTGATCCA 714
Db 1598 GAGAGGGGGTTTCGCCATGTTGACCGAGGTGGTCTCGAACTCTTGACGTCAAGTGATCCA 1157
Qy 715 CCTCCGTAGCTCTCAAAGTGTCTGGAATTTACAGGAAATTAATCCCAAGACATCAATCA 774
Db 1158 CCTCCGTAGCTCTCAAAGTGTCTGGAATTTACAGGAAATTAATCCCAAGACATCAATCA 1217
Qy 775 ATCTTTAAAGTACGGGTTTTCCTGATGATCAGTATACCTGAGCCAGACATGGAA 834
Db 1218 ATCTTTAAAGTACGGGTTTTCCTGATGATCAGTATACCTGAGCCAGACATGGAA 1277
Qy 835 ACTCTTGAGGAAAAGTTCTCAGATGTTTCATCTCTGTGGCTTGAATCTTCAAGGGGTGT 894
Db 1278 ACTCTTGAGGAAAAGTTCTCAGATGTTTCATCTCTGTGGCTTGAATCTTCAAGGGGTGT 1337
Qy 895 CTGAGATGATTCAGATGACAGATTAACTGTTCCCACTCTCGAGAGCTCTTACTTTT 954
Db 1338 CTGAGATGATTCAGATGACAGATTAACTGTTCCCACTCTCGAGAGCTCTTACTTTT 1397
Qy 955 GATTCTTTCAAGAGGCCCAAAATTAAGAGAAAGACAGCTTAATGAAGAAAGAAACAGAGA 1014
Db 1398 GATTCTTTCAAGAGGCCCAAAATTAAGAGAAAGACAGCTTAATGAAGAAAGAAACAGAGA 1457
Qy 1015 CGCCAAACAGT 1025
Db 1458 CGCCAAACAGT 1468

RESULT 8
AX286069
LOCUS AX286069 687 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 3 from Patent WO0179488.
ACCESSION AX286069
VERSION AX286069.1 GI:17045995
KEYWORDS Homo sapiens (human)
SOURCE
```



```

RESULT 10
AX056404
LOCUS AX056404 911 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 48 from Patent WO0073469.
ACCESSION AX056404
VERSION AX056404.1 GI:12229111
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1.
AUTHORS Plowman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
TITLE Protein kinases
JOURNAL Patent: WO 0373469-A 48 07-DEC-2000;
KEYWORDS Sugen, Inc. (US)
FEATURES
    source
        1..911
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT 293 a 182 c 178 g 258 t
ORIGIN
Query Match 51.6%; Score 537; DB 6; Length 911;
Best Local Similarity 87.0%; Pred. No. 7.2e-120;
Matches 643; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 285 AAATGGAGTTGCTGATGAGTGTCAAAAGGTTATATGGCAACACTTCAAGCTCTTAA 344
DB 1 AAATGGAGTTGCTGATGAGTGTCAAAAGGTTATATGGCAACACTTCAAGCTCTTAA 60

QY 345 TTCTGTCTATACATACTGTATTCACAGAGATATAAACCTCGAAATATTTCTATAAC 404
DB 61 TTCTGTCTATACATACTGTATTCACAGAGATATAAACCTCGAAATATTTCTATAAC 220

QY 405 TAAGCAAGGAATATCAAGATTGTGACTTCGGGTTTGCAAAATTCGTATCCAGGAGA 464
DB 121 TAAGCAAGGAATATCAAGATTGTGACTTCGGGTTTGCAAAATTCGTATCCAGGAGA 180

QY 465 TGCCTACACCGATTATGTAGTACAGAGTGTACCGAGCTCTGAACTTCTGTGGGAGA 524
DB 181 TGCCTACACCGATTATGTAGTACAGAGTGTACCGAGCTCTGAACTTCTGTGGGAGA 240

QY 525 TACTCAGTATGGTCTTCAGTCGATATATGGGCTATTGGTGTGTTTTCGAGAGCTCT 584
DB 241 TACTCAGTATGGTCTTCAGTCGATATATGGGCTATTGGTGTGTTTTCGAGAGCTCT 300

QY 585 GACGGCCAGCCACTGTGGCTGGAAATCAGATGTGACCAACTTATCTGATATACAG 644
DB 301 GACGGCCAGCCACTGTGGCTGGAAATCAGATGTGACCAACTTATCTGATATACAG 360

QY 645 AACCATAGTAGAGACGGGTTTTCGSCATSTTGACAGCGCTGGTCTCGAACTCTTGACGTC 704
DB 361 AACCATAGTAGAGACGGGTTTTCGSCATSTTGACAGCGCTGGTCTCGAACTCTTGACGTC 366

QY 705 AAGTGATCGACCTGCGGTAGCTCTCAAAAGTCTGGAATATACAGGAATTAATCCCAAG 764
DB 367 AAGTGATCGACCTGCGGTAGCTCTCAAAAGTCTGGAATATACAGGAATTAATCCCAAG 384

QY 765 ACATCAATCAATCTTTAAAGTAACGGGTTTTCATCGGCATCAGTATACCTGAGCCAGA 824
DB 385 ACATCAATCAATCTTTAAAGTAACGGGTTTTCATCGGCATCAGTATACCTGAGCCAGA 444

QY 825 AGACATGGAACTCTTGAGGAAAGTCTCAGATGTTTCATCTGTGGCTCTGAACTTCAT 884
DB 445 AGACATGGAACTCTTGAGGAAAGTCTCAGATGTTTCATCTGTGGCTCTGAACTTCAT 504

QY 885 GAAGGGGTCTCTGAGATGAAATCCAGATGACAGATTAACCTGTTCCCACTCTCGGAGAG 944
DB 505 GAAGGGGTCTCTGAGATGAAATCCAGATGACAGATTAACCTGTTCCCACTCTCGGAGAG 564

QY 945 CTCCTACTTTGATTTCTTTCAAGAGCGCCAAATTAAGAGAAAGACGCTATATCAAGAG 1004
DB 565 CTCCTACTTTGATTTCTTTCAAGAGCGCCAAATTAAGAGAAAGACGCTATATCAAGAG 624

QY 1005 AACACAGAGACGCGCAACAG 1023
DB 625 AACACAGAGACGCGCAACAG 643

RESULT 11
AX107712
LOCUS AX107712 561 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1 from Patent WO0123579.
ACCESSION AX107712
VERSION AX107712.1 GI:13923197
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1.
AUTHORS Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 1 05-APR-2001;
KEYWORDS Lexicon Genetics Incorporated (US)
FEATURES
    source
        1..561
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT 189 a 87 c 112 g 173 t
ORIGIN
Query Match 43.7%; Score 454.4; DB 6; Length 561;
Best Local Similarity 99.8%; Pred. No. 8.3e-100;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAAAGTATGAAAATTAGCTAGACTGGAGAGGGTCTTATCGGGTTGTATTCAA 60
DB 1 ATGAAAAGTATGAAAATTAGCTAGACTGGAGAGGGTCTTATCGGGTTGTATTCAA 60

QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTCTGTAAAAAATTTGTGNAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGTCTGTAAAAAATTTGTGNAATCTGAAGAT 120

QY 121 GATCCTGTGTTGTTAAGAAATAGCACCTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
DB 121 GATCCTGTGTTGTTAAGAAATAGCACCTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180

QY 201 CCAATCTTGTGAACCTCATCGAGGTCTCAGGAGAAAGGAAATGCAATTTAGTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTCTCAGGAGAAAGGAAATGCAATTTAGTTT 240

QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAACCCAAATGGAGTTGCTGAT 300

QY 301 GGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAATTTCTGTCATATACAT 360

QY 361 AACTGTATTCACAGATATATAAACCTGAAATATTTCTAATTAAGCAAGGAATAATC 420
DB 361 AACTGTATTCACAGATATATAAACCTGAAATATTTCTAATTAAGCAAGGAATAATC 420

QY 421 AAGATTTGTGACTTCGGGTTTGCAAAATTTCTGATT 456
DB 421 AAGATTTGTGACTTCGGGTTTGCAAAATTTCTGATT 456

RESULT 12
```

AX107718
LOCUS AX107718 594 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO023579.
ACCESSION AX107718
VERSION AX107718.1 GI:13923200
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.
Human kinase proteins and polynucleotides encoding the same
Patent: WO 0123579-A 7 05-APR-2001;
Lexicon Genetics Incorporated (US);
FEATURES
source
1..594
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 203 a 115 g 182 t
ORIGIN
Query Match 43.7%; Score 454.4; DB 6; Length 594;
Best Local Similarity 99.8%; Pred. No. 8.3e-100;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAAAGTATGAAAAATAGCTAAGACTGAGAGGGTCTTATGGGTGTATTCAAA 60
DB 1 ATGGAAAGTATGAAAAATAGCTAAGACTGAGAGGGTCTTATGGGTGTATTCAAA 60
QY 61 TCAGAAACAAACCTCTGGCAAGTAGTAGTGTATAAAATTTGTGGAACTCTGAAGAT 120
DB 61 TCAGAAACAAACCTCTGGCAAGTAGTAGTGTATAAAATTTGTGGAACTCTGAAGAT 120
QY 121 GATCTGTGTTTAAAGAAATAGCACTAGAGAAATACCTGTTCAGGCATTAACACAT 180
DB 121 GATCTGTGTTTAAAGAAATAGCACTAGAGAAATACCTGTTCAGGCATTAACACAT 180
QY 181 CCAATCTTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCAATTTAGTTTT 240
DB 181 CCAATCTTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCAATTTAGTTTT 240
QY 241 GAATAGTGTATCATACATCTTTTAATGCTGGAGAAACCCAAATGAGTTGCTGAT 300
DB 241 GAATAGTGTATCATACATCTTTTAATGCTGGAGAAACCCAAATGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
QY 361 AACTGTATTACAGAGATATAAACTGAAATATCTTAATTAACAGAGGAATATATC 420
DB 361 AACTGTATTACAGAGATATAAACTGAAATATCTTAATTAACAGAGGAATATATC 420
QY 421 AAGATTGTSACTTCGGGTTTCACAAATTCGTATT 456
DB 421 AAGATTGTSACTTCGGGTTTCACAAATTCGTATT 456
RESULT 13
LOCUS AX056405 2615 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 49 from Patent WO0073469.
ACCESSION AX056405
VERSION AX056405.1 GI:12229112
KEYWORDS Murinae gen. sp.
SOURCE
ORGANISM Murinae gen. sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
REFERENCE 1

AUTHORS Plowman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 49 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source
1..2615
/organism="Murinae gen. sp."
/mol_type="genomic DNA"
/db_xref="taxon:39108"
BASE COUNT 742 a 580 c 585 g 708 t
ORIGIN
Query Match 43.5%; Score 453.2; DB 6; Length 2615;
Best Local Similarity 73.1%; Pred. No. 1.6e-99;
Matches 658; Conservative 0; Mismatches 143; Indels 99; Gaps 2;
QY 135 GAAATAGCACTAAGAGAAATACGTNTGTTGAGCAATTAAGACATCCAAATCTTGAA 194
DB 1 GAAATAGCACTAAGAGAAATACGTNTGTTGAGCAATTAAGACATCCAAATCTTGAA 194
QY 195 COTCATCGAGGTGTTTCAGGAGAGAAAGGAAATGCAATTTAGTTTGAATCTGTGATCA 254
DB 50 COTCATCGAGGTGTTTCAGGAGAGAGAGAGATGCAATCTGTTTGAAGTCTGTGATCA 117
QY 255 TACACTTTTAATAGCTGGAAGAGAAACCCAAATGAGGTGCTGATGGAGTGATCAAAAG 314
DB 119 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGGAGTAAAG 177
QY 315 CGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACATACATCTGATTCACAG 374
DB 178 TGTGATGGAACACCTTCAAGCCCTTAATCTGTCACAGCAATTTGTTATCTGTCG 237
QY 375 AGATATAAAACCTGAAATATTTCTAATACTTAAGCAAGGAATTAATCAAGATTTGTGACTT 434
DB 238 CGATGTAACCTGAAACATCTCTAATAACCAAGAGGATGATAAAGATTTGTGACTT 297
QY 435 CGGTTTTGCACAAATCTGATTCAGAGAGATGCTACACCGATTAATCTAGCTAGAGATG 494
DB 298 TGGATTTGCACAAATCTGATTCAGAGAGAGAGCTACACAGACTATGTTGCCACCGGTG 357
QY 495 GTACCGAGCTCTGAACTCTTGTGGAGATACATCAAGTATGGTTCTTCAGTCGATATATG 554
DB 358 GTACCGAGCTCTGAACTCTTGTGGAGAGAGAGTACAGTCTCTCTGAGACGTGTG 417
QY 555 GCTATTGTTGTTGTTTTCAGAGAGCTCTGACAGGCGAGCCACTGTGGCTGGAAATC 614
DB 418 GGCCGTGCGCTGTGTTTGTGAGAGCTCTGACGGGTGAGCCACTCTGGCCGGGAAATC 477
QY 615 AGATGTGACCAACTTTATCTGATAATCAGAACTAGTAGAGACGGGTTTCGCCATGT 674
DB 478 CGAGCTGGACAGCTTTTACCTGATCATCGAGCGTTG----- 514
QY 675 TGACCGAGCTGGTCTCGAACTCTTTGACGTCAAGTGATCCACCTGCGCTAGCCTCTCAAG 734
DB 515 ----- 514
QY 735 TGCTGGAATTACAGGAAATTAATCCAGACATCAATCAATCTTTAAAGTAACGGTT 794
DB 515 -----GGGAAGCTGATTCAGAGACCACTCTATCTTTAGGAGTAACCGTT 561
QY 795 TTTCCATGGCATCAGTATACCTGAGCCAGAGACATCGAAACTCTTGAGGAAAGTTTCTC 854
DB 562 TTTCCGGGATCAGCATACCTGAAACAGAGGACATCGAGACTCTTGAAGAAAATTTCTC 621
QY 855 AGATCTTCATCTGTGCTCTGAATCTTCATGAAGGGGTGCTGAAGATGAATCCAGATGA 914
DB 622 AAATGTTTCAGCTGTGGCTTTAAGTTTCATGAAGGGGATGCTGAAGATGAATCTCTGATGA 681
QY 915 CAGATTAACTGTTCCCAACTCTCTGGAGAGCTCCCTACTTTGATTTCTTTCAAGAGGCCCA 974
DB 682 GAGGCTGACCTGTGCCAGCTGCTGGACAGTCTACTTTGAGTCTTTTCAAGAGGATCA 741
QY 975 AATTAAAGAAAGCACGCTAATGAAGGAGAAACAGAAAGACGCCAACAGGTACTTCCGCT 1034

Query Match: 33.0% Score 343.8 DB 6: Length 1612:

```
Best Local Similarity 71.2%; Pred. No. 5.9e-73;
Matches 469; Conservative 0; Mismatches 187; Indels 3; Gaps 1;

QY 1 ATGGAAAGATATGAAATAATAGCTAAGACTGGAGAGGCTCTTATGGGTTGTTATTCAAA 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
466 ATGGAGAAATATGAAATAATGGGAAATGGAGAGGATCCTATGGAGTTGTTTCAA 525
QY 61 TGCAGAAACAAACCTCTGGCAAGTAGTAGTGTATAAAATTTTGGAACTCTGAAGAT 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
526 TGTAGAAACAGGACACGGGTAGATTGTGCCATCAGAAAGTTCTGGATCAGAAAGAT 585
QY 121 GATCTCTGTTTAAAGAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACAT 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
586 GACCCGTGTCAAAAGAAATTTGCCCTTCGGAAATCCGAATGCTCAAGCAACTCAAGCAT 645
QY 181 CCAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTAAGTTTTT 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
646 CCAACCTTGTTAACTCTCTGGAAGTCTTCAGGAGGAAACGGAGGCTTCACCTGTGTTT 705
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
706 GAATATTGTGACCAACACAGTCTCCATGAGTTGGACAGATACCAAGAGGGGTACCAAA 765
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
766 CATCTCGTGAGAGGATCACTTGGCAGACACTGCCAAGCTGTAAATTTTGCATTAACAC 825
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATCTTAATACTAAAGCAAGGAATATC 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
826 AATTGCATACATAGAGAGCTGAAGCCAGAAATACTCTCATCAGCAACATTCCTGTGATT 885
QY 421 AAGATTGTGACTTCGGGTTTGCACAAATTCGTATT---CCAGGAGATGCTACACCGAT 477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
886 AAGCTTTGTGACTTTGGATTTCCTCGGCTTTTGACTGGACCGAGTGACTACTATACAGAC 945
QY 478 TATGAGCTACGAGATGATACCGAGCTCCTGAACCTTCTGTGGAGATACCTCAGTATGTT 537
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
946 TACGTGGCTACCAAGTGTGTACCGCTCCCTGAGCTGTGTGGGGGACACGCACTACGGC 1005
QY 538 TCTTCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCCTGACAGGCCAGCCA 597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1006 CCCCCGGTGGATGTTTGGGCAATGGCTGTGTCTTTGCTGAGCTGCTGTACAGGAGTGCCT 1065
QY 598 CTGTGCGCTGGAAATCAGATCTGGACCAACTTTATCTGATAATCAGAACACTAGTAGA 656
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1066 CTGTGCGCCAGGAAATCGATGTGATCAGCTGTATCTGATTAGGAACACCTTGGGGA 1124
```

Search completed: November 15, 2003, 19:04:18
Job time : 4087.56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 13, 2003, 14:12:21 ; Search time 39.8369 Seconds
(without alignments)
1382.592 Million cell updates/sec

Title: US-09-671-050-10

Perfect score: 1820

Sequence: 1 MEYKELAKTGEISYGVVFK.....RKARNEGRRRQVVPLKS 347

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1:07863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1:07863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1820	100.0	347	22	AAE00494 Human kinase #5.
2	1796	98.7	356	22	AAE00491 Human kinase #2.
3	1636	89.9	315	22	AAE00495 Human kinase #6.
4	1613	88.6	342	22	AAE19152 Human kinase polyp
5	1612	88.6	324	22	AAE00492 Human kinase #3.
6	1593	87.5	360	22	AAU03525 Human protein kinase
7	1270.5	69.8	296	22	AAE65643 Novel protein kinase
8	1162.5	63.9	358	23	ABP62954 Human polypeptide
9	1134	62.3	228	23	AAG78547 Human kinase 14257

10	1123	61.7	247	22	AAE65642 Novel protein kinase
11	1001.5	55.0	392	22	ABE63118 Drosophila melanog
12	948	52.1	197	24	ABU11689 Human MDT polypep
13	941.5	51.7	566	21	RAY90724 Rabbit KKIARE kin
14	911.5	50.1	362	24	ABP96087 Human protein kinase
15	791	43.5	187	22	AAE00490 Human kinase #1.
16	791	43.5	198	22	AAE00493 Human kinase #4.
17	758.5	41.7	455	21	AAU70126 Human lost in leuk
18	758.5	41.7	591	22	AAU03524 Human protein kinase
19	737.5	40.5	457	21	AAU70125 Rat lost in leukae
20	737.5	40.5	505	21	AAU70124 Rat lost in leukae
21	712.5	38.1	154	24	ABU05290 Human diagnostics
22	666	36.6	534	22	AAE65645 Novel protein kinase
23	569	31.3	224	22	ABG06142 Novel human diagno
24	565	31.0	333	22	AAU39276 Human polypeptide
25	563	30.9	544	19	AAW85028 CDK2-green fluore
26	558	30.7	352	22	AAU41062 Human polypeptide
27	557	30.6	298	20	AAW95689 Homo sapiens Cdc2
28	557	30.6	544	19	AAW85029 Green fluorescent
29	555	30.5	297	23	ABW79004 Human 10ckshsl pro
30	551.5	30.3	294	20	AAW95690 Oryza sativa Cdc2
31	551.5	30.3	294	21	AAG11214 Arabidopsis thalia
32	551.5	30.3	294	21	AAG54042 Arabidopsis thalia
33	551.5	30.3	350	21	AAG11213 Arabidopsis thalia
34	547.5	30.1	270	20	AAU43932 Human protein kinase
35	546.5	30.0	294	21	AAU40809 Zea mays protein f
36	545	29.9	297	20	AAW95688 Homo sapiens Cdc2
37	545	29.9	297	22	AAAG7434 Amino acid sequenc
38	545	29.9	297	22	AAAG7613 Amino acid sequenc
39	545	29.9	297	24	ABR47405 Breast cancer asso
40	545	29.9	297	24	AAE34491 Human CDC2 protein
41	545	29.9	297	24	ABU56522 Lung cancer-associ
42	541.5	29.8	294	21	AAU33356 Zea mays protein f
43	541	29.7	297	20	AAW95687 Rattus norvegicus
44	541	29.7	298	21	AAU52184 Cyclin-dependent k
45	540.5	29.7	297	22	ABE61462 Drosophila melanog

ALIGNMENTS

RESULT 1
AE00494
10 AAE00494 standard; Protein; 347 AA.
XX
AC AAE00494;
XX
19-JUN-2001 (first entry)
DE Human kinase #5.
XX
Human; kinase; gene therapy; bio-reactor; mental disorder;
XX biological disorder.
OS Homo sapiens.
XX
XX WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Dorohc G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX WPI: 2001-266166/27.
XX DR N-PSDB; AAD03816.
XX
XX New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for
PT

compounds useful for treating mental, biological or medical diseases -
Claim 2; Page 32-33; 39pp; English.

The present sequence is novel human protein (NHP) known as human kinase. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.

Sequence 347 AA;

Query Match 100.0%; Score 1820; DB 22; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.7e-186;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEISGVVFKCRNKTSGQVAVKFESEDDPVVKIALREIRMLKQLKH 60
DB 1 NEKYEKLAKTGEISGVVFKCRNKTSGQVAVKFESEDDPVVKIALREIRMLKQLKH 60
QY 61 PNLVNLLEVFPRKRMHLVFYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCCHI 120
DB 61 PNLVNLLEVFPRKRMHLVFYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCCHI 120
QY 121 NCIRHDIKPNILITKQGIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180
DB 121 NCIRHDIKPNILITKQGIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPPA 240
DB 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPPA 240
QY 241 VASQSAGITGKLIIPRHQSIFKSNCFPHGISIPEPEDMETLEEFSDVHPVALNFMKGLK 300
DB 241 VASQSAGITGKLIIPRHQSIFKSNCFPHGISIPEPEDMETLEEFSDVHPVALNFMKGLK 300
QY 301 MNPDRLTCSQLLESSYFDSFCEAQIKRKARNEGRRRROOVFLKS 347
DB 301 MNPDRLTCSQLLESSYFDSFCEAQIKRKARNEGRRRROOVFLKS 347

RESULT 2
AAE00491
ID AAE00491 standard; Protein; 356 AA.

XX AAE00491;

AC AAE00491;

XX 19-JUN-2001 (first entry)

XX Human Kinase #2.

XX Human; kinase; gene therapy; bioreactor; mental disorder;

XX Biological disorder.

XX Homo sapiens.

OS

XX WO200123579-A1.
XX CS-APR-2001.
XX 27-SEP-2000; 2000WO-US26621.
XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Turner CA, Nehls X, Friedrich G, Zambrowicz B, Sands AT;
XX MPI; 2001-266.66/27.
XX N-PSDB; AND03813.
XX New isolated human kinase polynucleotide useful for generating for
XX antibodies, as reagents in diagnostic assays and for screening for
XX compounds useful for treating mental, biological or medical diseases -
XX Disclosure; Page 28-29; 38pp; English.
XX The present sequence is novel human protein
XX (NHP) known as human kinase. The human kinases share structural
XX similarity with animal kinases, more particularly serine or
XX threonine protein kinases. Human kinase cDNA is useful for the
XX detection of mutant human kinase for the diagnosis of disease,
XX and also as a therapeutic. It is useful for screening drugs
XX effective in the treatment of symptomatic or phenotypic
XX manifestations perturbing the normal function of NHP in the
XX body. The NHP nucleotide sequences are useful for generation of
XX antibodies, as reagents in diagnostic assays, for the
XX identification of other cellular gene products related to human
XX kinases, and as reagents in assays for screening compounds that
XX are useful for treating mental, biological or medical disorders.
XX NHP oligonucleotides are used as probes. The labelled NHP probes
XX are useful for screening human genomic library for identifying
XX polymorphisms and as primers in amplification assays to detect
XX mutations within the exons, introns and splice sites that can
XX be used in diagnostics and pharmacogenomics. Nucleotide construct
XX encoding NHP products are used to genetically engineer cells
XX in vivo that functions as bioreactors in the body delivering a
XX continuous supply of NHP to the body. Nucleotide constructs
XX encoding functional NHPs are used in gene therapy for the
XX modulation of NHP expression.

Sequence 356 AA;

Query Match 98.7%; Score 1796; DB 22; Length 356;
Best Local Similarity 100.0%; Pred. No. 6.6e-184;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEISGVVFKCRNKTSGQVAVKFESEDDPVVKIALREIRMLKQLKH 60
DB 1 NEKYEKLAKTGEISGVVFKCRNKTSGQVAVKFESEDDPVVKIALREIRMLKQLKH 60
QY 61 PNLVNLLEVFPRKRMHLVFYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCCHI 120
DB 61 PNLVNLLEVFPRKRMHLVFYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCCHI 120
QY 121 NCIRHDIKPNILITKQGIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180
DB 121 NCIRHDIKPNILITKQGIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPPA 240
DB 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPPA 240
QY 241 VASQSAGITGKLIIPRHQSIFKSNCFPHGISIPEPEDMETLEEFSDVHPVALNFMKGLK 300
DB 241 VASQSAGITGKLIIPRHQSIFKSNCFPHGISIPEPEDMETLEEFSDVHPVALNFMKGLK 300
QY 301 MNPDRLTCSQLLESSYFDSFCEAQIKRKARNEGRRRROOV 342

DB 30: XNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRQV 342

RESULT 3

AAE00495
ID AAE00495 standard; Protein; 315 AA.

AC AAE00495;

XX 19-JUN-2001 (first entry)

XX Homo sapiens

XX Human kinase #6.

XX Human; Kinase; Gene therapy; bioreactor; mental disorder;
KW biological disorder.

XX Homo sapiens.

XX WO200123579-A1.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26621.

XX 28-SEP-1999; 99US-015451.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX WPI; 2001-266166/27.

XX N-PSDB; AAD03817.

XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX
XX Claim 3; Page 33-34; 38pp; English.

CC The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 315 AA;

Query Match 89.9%; Score 1636; DB 22; Length 315;
Best Local Similarity 90.8%; Pred. No. 8.6e-167;
Matches 315; Conservative 0; Mismatches 6; Indels 32; Gaps 1;

QY 1 MEKYEKANTGEGSYGVWFKCRNKTSQGVVAVKFESEDDPVVKKIALREIRMLKQKH 60

DB 1 MEKYEKANTGEGSYGVWFKCRNKTSQGVVAVKFESEDDPVVKKIALREIRMLKQKH 60

QY 61 PNLVNLTIEVFRKRKMEVFVEYCDHTLLNELRNPNGVADGVKSVLWOTLQALNFCIH 120
DB 61 PNLVNLTIEVFRKRKMEVFVEYCDHTLLNELRNPNGVADGVKSVLWOTLQALNFCIH 120
QY 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDVATRWYRAPELLVGDTOYGS 180
DB 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDVATRWYRAPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240
DB 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240
QY 241 VASQSAGITKGLIPRHQSIFKSNFGFFHGISIPEDMETLEEFSDVHPVALNFMKGLCK 300
DB 241 VASQSAGITKGLIPRHQSIFKSNFGFFHGISIPEDMETLEEFSDVHPVALNFMKGLCK 300
QY 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRQVLPKLS 347
DB 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRQVLPKLS 347
QY 269 MKPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRQVLPKLS 315
DB 269 MKPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRQVLPKLS 315

RESULT 4

AAE19152

ID AAE19152 standard; Protein; 342 AA.

XX AAE19152;

XX 21-MAY-2002 (first entry)

XX Human kinase polypeptide (PKIN-10).

XX Human; kinase polypeptide; PKIN-10; gene therapy; Addison's disease;
KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
KW cholestasis; cardiast; cardiovascular disorder; Niemann-Pick's disease;
KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
KW drug screening; transgenic animal; antinflammatory; hepatotropic;
KW hypotensive; anti-HIV; enzyme.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 4...286 /note= "Eukaryotic protein: kinase domain"

XX WO200208399-A2.

XX 31-JAN-2002.

XX 20-JUL-2001; 2001WO-US21092.

XX 21-JUL-2000; 2000US-220038P.

XX 28-JUL-2000; 2000US-222112P.

XX 04-AUG-2000; 2000US-222831P.

XX 11-AUG-2000; 2000US-224729P.

XX (INCY-) INCYTE GENOMICS INC.

XX (THOR-) THORNTON M.

XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Wallia NK;
PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P;
PI Ding L, Yao MG, Elliott VS, Resipon SA, Kearney L, Lu DAM;
PI Greenwald SR, Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J;
PI Hillman JL;

XX WPI; 2002-206083/26.

XX N-PSDB; AAD30557.

XX New human kinase polypeptide, useful in diagnosis, prevention and

PT treatment of cancer, immune disorder, growth and developmental
PT disorder, cardiovascular disorder and lipid disorder -

PS Claim 1: Page 147-148; 196pp; English.

XX The present invention relates to an isolated human kinase polypeptide
CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
CC useful for diagnosing, treating and preventing cancer (e.g., leukemia,
CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
CC drug screening techniques and to analyse the proteome of a tissue or cell
CC type. PKIN is useful for creating knockin humanised animals or transgenic
CC animals to model human diseases, in somatic or germline gene therapy, to
CC generate a transcript image of a tissue or cell type, for detecting
CC differences in the chromosomal location due to translocation, inversion,
CC etc., among normal, carrier or affected individuals, and as hybridisation
CC probes for mapping naturally occurring genomic sequences. PKIN is useful
CC in southern or northern analysis, dot blot or other membrane-based
CC technologies, in PCR technologies, in dipstick, pin, multiforamat enzyme
CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
CC fluids or tissues from patients to detect altered PKIN expression. The
XX present sequence is human PKIN-10.

XX Sequence 342 AA;

Query Match 88.6%; Score 1613; DB 23; Length 342;

Best Local Similarity 89.94; Pred. No. 2.9e-164;

Matches 312; Conservative 1; Mismatches 0; Indels 34; Gaps 2;

Qy 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALREIRMLKQLKH 60

Qy 61 PNLVNLIEVFRRRKXKHLVFEYCDHTLLNLELRNPNGVADGVKSVLMQTLQALNFCCHI 120

Db 61 PNLVNLIEVFRRRKXKHLVFEYCDHTLLNLELRNPNGVADGVKSVLMQTLQALNFCCHI 120

Qy 121 NCIHRLDKPENILITKGGIIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180

Db 121 NCIHRLDKPENILITKGGIIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180

Qy 181 SVDIWAIGCVFAELLTGQPLMPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

Db 181 SVDIWAIGCVFAELLTGQPLMPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

Qy 241 VASOSAGITCKLIPRHOSIFKSNFGFHGISIPEDMETLEEKFSVDHPVAFNFKGCLK 300

Db 241 VASOSAGITCKLIPRHOSIFKSNFGFHGISIPEDMETLEEKFSVDHPVAFNFKGCLK 300

Qy 218 -----GKLIIPRHOSIFKSNFGFHGISIPEDMETLEEKFSVDHPVAFNFKGCLK 268

Db 269 MNPDDRUTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQV -GVLPJ 345

Qy 301 MNPDDRUTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQV -GVLPJ 345

Db 269 MNPDDRUTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQV -GVLPJ 315

RESULT 5

AAE00492

ID AAE00492 standard; Protein: 324 AA.

XX AC AAE00492;

XX AC AAE00492;

DT 19-JUN-2001 (first entry)

XX Human kinase #3.

XX Human; kinase; gene therapy; bioreactor; mental disorder;

XX biological disorder.

XX Homo sapiens.

XX WC200123579-A1.

XX PD 05-APR-2001.

XX 27-SEP-2000; 2000WC-US26621.

XX 28-SEP-1999; 99US-C156511.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Kehls M, Friedrich G, Zambrowicz B, Sands AT;

XX WPI; 2001-266166/27.

XX N-PSDB; AA030814.

XX New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for
XX compounds useful for treating mental, biological or medical diseases -

XX Disclosure; Page 30; 38pp; English.

XX The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 324 AA;

Query Match 88.6%; Score 1612; DB 22; Length 324;

Best Local Similarity 90.6%; Pred. No. 3.4e-164;

Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

Qy 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALREIRMLKQLKH 60

Qy 61 PNLVNLIEVFRRRKXKHLVFEYCDHTLLNLELRNPNGVADGVKSVLMQTLQALNFCCHI 120

Db 61 PNLVNLIEVFRRRKXKHLVFEYCDHTLLNLELRNPNGVADGVKSVLMQTLQALNFCCHI 120

Qy 121 NCIHRLDKPENILITKGGIIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180

Db 121 NCIHRLDKPENILITKGGIIKICDFGFAQLIPGDAYTDVATRWYRAPPELLVGDTOYGS 180

Qy 181 SVDIWAIGCVFAELLTGQPLMPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

Db 181 SVDIWAIGCVFAELLTGQPLMPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

Qy 241 VASOSAGITCKLIPRHOSIFKSNFGFHGISIPEDMETLEEKFSVDHPVAFNFKGCLK 300

Db 241 VASOSAGITCKLIPRHOSIFKSNFGFHGISIPEDMETLEEKFSVDHPVAFNFKGCLK 300

Qy 218 -----GKLIIPRHOSIFKSNFGFHGISIPEDMETLEEKFSVDHPVAFNFKGCLK 268

Qy 301 MNPDDRUTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQV 342

CC activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative-stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders.

XX
SQ Sequence 296 AA;

Query Match 69.8%; Score 1270.5; DB 22; Length 296;
Best Local Similarity 80.5%; Pred. No. 1.5e-127;
Matches 243; Conservative 19; Mismatches 5; Indels 35; Gaps 3;

QY 46 KIALEIRMLKQLKHNLVLEIFRRKRKMLHVFYCDHTLLNELRNNGVADGVKS 105
DB : KIALEIRMLK-LKHPNLVNLVLEIFRRKRKMLHVFYCDHTLLNELRNNGVADGVKS 59

QY 106 VLWQTQALNFCIHNCIHRDIXPENILITKQGIKICDFGPAQLIPGDAYTYVATEW 165
DB 60 VLWQTQALNFCIHNCIHRDIXPENILITKQGIKICDFGPAQLIPGDAYTYVATEW 119

QY 166 YRAPPELLVGDYQYGVSSVDVWAVGCVFAELLTGQPLWPKKSDVDQYLIIRLTVEGTFRHV 225
DB 120 YRAPPELLVGDYQYGVSSVDVWAVGCVFAELLTGQPLWPKKSDVDQYLIIRLTVEGTFRHV 171

QY 226 DQAGLELLTSSDPPAVASQASAGITKLPFHQSIFKNGFFHGSIPEPEDMETLSEKFS 285
DB 172 -----GKLIPIHQSIKFRNQCFRFGSIPEPEDMETLSEKFS 207

QY 286 DVHPVALNFMKGCLKQNPDDRLTSCQLLESYFSPQEAQIKRKARNEGRRRQ--QVL 343
DB 208 NVQPVALSFMKGCLKQNPDDRLTCAQLDSAYFSPQEAQIKRKARNEGRRRQ--QVL 267

QY 344 PL 345
DB 268 PL 269

RESULT 8
ID ABP62954 standard; Protein; 358 AA.
XX
AC ABP62954;
XX
XX
DT 14-OCT-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 391.
XX
XX Human; vulnery; dermatological; neuroprotective; nootropic; cancer; anti-parkinsonian; immunostimulant; cytostatic; immunosuppressive; antidiabetic; antiallergic; gene therapy; wound healing; tissue repair; burn; central nervous system disorder; Alzheimer's disease; Parkinson's disease; Huntington's disease; immune disorder; autoimmune disorder; multiple sclerosis; diabetes; a1e1g1.

OS Homo sapiens.
XX
XX W20020218424-A2.
PN
XX
PE 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US27093.
XX
XX 01-SEP-2000; 2000US-0654935.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang C;
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX
XX WPI: 2002-583321/62.
DR N-PSDB; ABQ93433.

XX New polynucleotide and polypeptides, useful for treatment and diagnosis of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple sclerosis, diabetes and allergies.

XX
PS Claim 20: SEQ ID NO 391; 284pp + Sequence Listing; English.

XX
XX The invention relates to an isolated polynucleotide (I) comprising one of 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising administering to a mammalian subject a composition comprising the protein (I) encoded by (I) (ABP62809-ABP63053) or an antibody (II) to (I). (II), (I) and (III) are useful for diagnostic evaluation of disorders. (I) is useful for gene therapy of diseases and (II) can be used for therapeutic treatment. Diseases that may be treated include wound healing and tissue repair, burns, central nervous system disorders (e.g. Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple sclerosis, diabetes and allergies.

CC
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 358 AA;

Query Match 63.9%; Score 1162.5; DB 23; Length 358;
Best Local Similarity 63.8%; Pred. No. 8.3e-116;
Matches 213; Conservative 42; Mismatches 46; Indels 33; Gaps 2;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGOVAVKVFESDDPVVKIALREIRMLKQLKH 60
DB 2 MEKYEKIGKIGESYGVVFKCRNRDTQGVIAIKKFLSEDDPVVKIALREIRMLKQLKH 61

QY 61 PNLVNLIEVFRKRKMLHVFYCDHTLLNELRNNGVADGVKSVLWQTQALNFCIH 120
DB 62 PNLVNLIEVFRKRKMLHVFYCDHTLVHLDYQGVPEHVKVKSITWQTQAVNFCIKH 121

QY 121 NCIHRDIXPENILITKQGIKICDFGPAQLIPGDAYTYVATEWYRAPPELLVGDYQY 179
DB 122 NCIHRDIXPENILITKQGIKICDFGPAQLIPGDAYTYVATEWYRAPPELLVGDYQY 181

QY 180 SSVDIWAIGCVFAELLTGQPLWPKKSDVDQYLIIRLTVEGTFRHVDPQAGLELLTSSDPP 239
DB 182 PPVDVWALGCVFAELLTGQPLWPKKSDVDQYLIIRLTVEGTFRHVDPQAGLELLTSSDPP 219

QY 240 AVASQASAGITKLPFHQSIFKNGFFHGSIPEPEDMETLSEKFSQVDPVVALNFMKGCL 299
DB 220 -----GDLIPRHOQVFNSTNQYFSGVKIPEDMEDPELKLFPNISYPALGLLKGCL 269

QY 300 KMNPPDRLTCSQLLESYFSPQEAQIKRKARNE 333
DB 270 HMDPTERTLTCQLLHPYFENIREIDLAKEHNK 303

RESULT 9
AAG78547
ID AAG78547 standard; Protein; 228 AA.
XX
AC AAG78547;
XX
DT 05-MAR-2002 (first entry)
XX
XX Human kinase 14257 amino acid sequence.

XX
XX Protein kinase; enzyme; cytosolic; osteopathic; hepatotropic; antidiabetic; neuroprotective; antiarthritic; dermatological; immunosuppressive; antiinflammatory; antithyroid; antipsoriatic; ophthalmological; antiallergic; antiasthmatic; antiatherosclerotic; hypertensive; vasotropic; antiarrhythmic; virucide; anorectic; metabolic; immunomodulator; analgesic; cellular proliferative disorder; cancer; acute lymphoblastic leukaemia; Hodgkin's disease; bone metabolism disorder; osteoporosis; immune system disorder; inflammatory; diabetes mellitus; osteoarthritis; asthma;

KW cardiovascular disorder; hypertension; coronary artery disease;
KW endothelial cell disorder; psoriasis.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Domain 4..218
FT FT /note= "eukaryotic protein kinase domain"
FT Modified-site 9..15
FT FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 23..28
FT FT /note= "N-glycosylation site"
FT Modified-site 27..32
FT FT /note= "N-myristoylation site"
FT Modified-site 38..41
FT FT /note= "casein kinase II phosphorylation site"
FT Modified-site 97..102
FT FT /note= "N-myristoylation site"
FT Modified-site 132..134
FT FT /note= "serine/threonine protein kinase active-site
FT FT signal site"
FT Domain 161..218
FT FT /note= "kinase transferase protein
FT FT serine/threonine-protein ATP-binding II phosphorylation
FT FT casein alpha chain domain"
FT Modified-site 189..183
FT FT /note= "casein kinase II phosphorylation site"
FT Modified-site 188..193
FT FT /note= "N-myristoylation site"
FT Modified-site 204..211
FT FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 205..208
FT FT /note= "casein kinase II phosphorylation site"
XX WD200179488-A2.
PN
XX
XX 25-OCT-2001.
XX
XX
XX 13-APR-2001; 2001WO-US12186.
XX
XX 13-APR-2000; 2000US-196910P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Kapeller-Libermann R;
PI
XX
XX WPI; 2002-034355/04.
XX N-PSDB; AAF64248.
XX
XX New 14257 polypeptides (protein kinases), useful as diagnostic targets
XX and therapeutic agents for controlling cellular proliferation and/or
XX differentiative disorder, bone disorders, immune disorders and
XX cardiovascular disorders
XX
XX Claim 1c; Fig 1A; 98pp; English.
XX
XX The invention relates to an isolated 14257 polypeptide and nucleic
XX acid encoding it. The 14257 protein is a protein kinase that acts as a
XX modulating agent in regulating a variety of cellular processes.
XX Including cell proliferation, differentiation, growth and division.
XX The activity of the protein of the invention may be described as;
XX cytostatic; osteopathic; hepatotropic; antidiabetic; neuroprotective;
XX antiarthritic; dermatological; immunosuppressive; antinflammatory;
XX antithyroid; antipsoriatic; ophthalmological; antiallergic;
XX antiasthmatic; antiatherosclerotic; hypotensive; vasotropic;
XX antiarrhythmic; virucide; anorectic; metabolic; immunomodulator and
XX analgesic. The protein of the invention may act as a novel diagnostic
XX target or therapeutic agent controlling certain disorders, for example
XX kinase-associated or other 14257-associated disorders. These may include
XX cellular proliferative disorders such as cancers e.g. acute lymphoblastic
XX leukemia or Hodgkin's disease. Other disorders include bone metabolism
XX disorders such as osteoporosis, disorders of the immune system, e.g.
XX inflammatory, diabetes mellitus, osteoarthritis and asthma. Proteins of

CC the invention may also be of use as therapeutic agents in cardiovascular
CC disorders such as hypertension and coronary artery disease, and some
CC endothelial cell disorders, including psoriasis. The current
CC sequence represents a human kinase 14257 amino acid sequence.
XX
SQ Sequence 228 AA;
Query Match 62.3%; Score 1134; DB 23; Length 228;
Best Local Similarity 98.2%; Pred. No. 4.7e-113;
Matches 213; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MEKYELAKTGEYSYGVVFKCRNKTSGQVAVKXVFSEDDPVVKIALREIRMLKQLKH 60
DB 1 MGKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKXVFSEDDPIVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKHLLVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCIH 120
DB 61 PNLVNLIEVFRKRKHLLVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCIH 120
QY 121 NC1HRDVKPENILITKQGIKICDPGFAQILIPGDAYTDYVATRWYRAPELLVGDYGS 180
DB 121 NC1HRDVKPENILITKQGIKICDPGFAQILIPGDAYTDYVATRWYRAPELLVGDYGS 180
QY 161 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIITL 217
DB 161 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIITL 217
RESULT 10
AAB65642
ID AAB65642 standard; Protein; 247 AA.
XX AAB65642;
XX
XX 27-XAR-2001 (first entry)
XX
XX Novel protein kinase, SEQ ID NO: 169.
XX
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
XX immunosuppressive; cardiant; renal; antinflammatory; antiasthmatic;
XX dermatological; antidiabetic; antifertility; gene therapy; vaccine;
XX immune disorder; cardiovascular disease; neurodegenerative disease;
XX cancer; autoimmune disorder; stroke; inflammatory bowel disease;
XX inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
XX Homo sapiens.
XX
XX WD200073469-A2.
XX
XX 07-DEC-2000.
XX
XX 26-MAY-2000; 2000WO-US14842.
XX
XX 28-MAY-1999; 99US-0136503.
XX
XX (SUGEN-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX WPI; 2001-032161/04.
XX N-PSDB; AAF44669.
XX
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
XX treating immune-related diseases and disorders, cardiovascular disease,
XX neurodegenerative diseases and/or cancers -
XX
XX Claim 10; Fig 1; 310pp; English.
XX
XX The present sequence is a novel protein kinase. The novel protein kinases
XX and the nucleic acids that encode them may be used in the treatment and
XX diagnosis of diseases associated with inappropriate kinase expression
XX such as immune-related diseases and disorders, cardiovascular disease,
XX neurodegenerative diseases and/or cancers. The nucleic acids and

CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.

XX Sequence 247 AA;
SQ Query Match 61.7%; Score 1123; DB 22; Length 247;
Best Local Similarity 86.1%; Pred. No. 8.1e-112;
Matches 217; Conservative 1; Mismatches 0; Indels 34; Gaps 2;
QY 96 NGVADGVKSVLWQTLQALNFCIHNCIHRDIPENILITKQGIKICDFGFAQILPGD 155
DB 1 NGVADGVKSVLWQTLQALNFCIHNCIHRDIPENILITKQGIKICDFGFAQILPGD 60
QY 156 AYTDVATRWYRAPPELLVGDYQSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIR 215
DB 61 AYTDVATRWYRAPPELLVGDYQSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIR 120
QY 216 TLVETGERHVDQAGLELLTSSDPPAVASQSAGITGKLIPIRHQSIFKNGFFHGSIPEPE 275
DB 121 TL-----GKLIPIRHQSIFKNGFFHGSIPEPE 148
QY 276 DMETLEEKFSVHPVVALNFMKGCGLMNPDDRLTCSQLLESSYFDSFGEA:KXKARNEGR 335
DB 149 DMETLEEKFSVHPVVALNFMKGCGLMNPDDRLTCSQLLESSYFDSFGEA:KXKARNEGR 208
QY 336 NRRRO--QVLEPL 345
DB 209 NRRROQQLPL 220

RESULT 11
ABB63118
ID ABB63118 standard; Protein: 392 AA.
AC ABB63118;
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 16146.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
PN WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEXE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
P: WPI; 2001-656860/75.
DR N-PSDB; ABL07221.
XX New isolated nucleic acid detection reagent for detecting 1300 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT Interactions -
XX Disclosure; SEQ ID NO 16146; 2lpp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC sequences (ABU01840-ABU16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 392 AA;
SQ Query Match 55.0%; Score 1001.5; DB 22; Length 392;
Best Local Similarity 53.0%; Pred. No. 1.9e-98;
Matches 184; Conservative 65; Mismatches 61; Indels 37; Gaps 3;
QY 1 MEKYELAKTGESSGYVFKCNKTSQGVAVKVFSEDDPVVKIALRIEIRMLKQKH 60
DB 1 MJRYEKLSRLGEGSYGVVYKCRDRETGALVAVKRFVEEDDPAIRKIALRIEIRMLKQKH 60
QY 61 PNLVNLIEVFRKRNKH:VFYECDHLLNELERNPVGADGVKSVLWQTLQALNFCIH 120
DB 61 PNLVNLIEVFRKRNHLVFEFCBLTVLHELHPQCPPEHLTKQICYQTLGVAYCHKQ 120
QY 121 NCIHRIKPEINILITKQGIKICDFGFAQILPGDVTYVATRWYRAPPELLVGDYQGS 180
DB 121 GCLHRDIPENILITKQGIKICDFGFAQILPGDVTYVATRWYRAPPELLVGDYQGS 180
QY 181 SYD:WAIGCVFAELLTGQPLWPKSDVDQLYLIRTLVETGERHVDQAGLELLTSSDPPA 240
DB 181 PVDVWAIGCLFAELVRGEALWPKSDVDQLYLIRTL----- 217
QY 241 VASQAGITGKLIPIRHQSIFKNGFFHGSIPEPEDMETLEEKF---SDVHPVALNFMKG 297
DB 218 -----GDLPIRHQIFGQNEVFKGITLPVPTLEPLEDKMPAKSQKNPLTIDFLKX 268
QY 298 CLMNPDDRLTCSQLLESSYFDSF--GEA:KXKARNEGRNRRQOV 342
DB 269 CLDKPTKRWSCKLTKHSYFDDYIAKORELHVNSLEAANLROOOL 315

RESULT 12
ABU11689
ID ABU11689 standard; Protein: 197 AA.
AC ABU11689;
XX 13-FEB-2003 (first entry)
XX Human YDDT polypeptide SEQ ID 636.
XX YDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-Hiv;
KW haemostatic; nephrotropic; antianaemic; antiproliferative; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis.
XX Homo sapiens.
XX OS
XX WC200279449-A2.
PN 10-OCT-2002.
PD 27-MAR-2002; 2002WO-US09944.
XX PF

XX 28-MAR-2001; 2001US-279619P.
 PR 29-MAR-2001; 2001US-280367E.
 PR 29-MAR-2001; 2001US-280368P.
 PR 16-MAY-2001; 2001US-291280P.
 PR 17-MAY-2001; 2001US-291829P.
 PR 17-MAY-2001; 2001US-291849P.
 PR 19-JUN-2001; 2001US-299428P.
 PR 20-JUN-2001; 2001US-299776P.
 PR 20-JUN-2001; 2001US-300001P.
 XX (INCYTE GENOMICS INC.
 PA Daffo A, Jones AB, Tran AB, Gietzen D, Chinn J;
 XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
 P Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gersin EH;
 P Peralta CH, David MH, Lewis SA, Chen AJ, Parzer SR, Harris B;
 P Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX WPI; 2003-058431/05.
 DR N-PSDB; ABX34679.
 DR
 XX New purified disease detection and treatment molecule proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing cancers
 P (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
 P or hepatitis -
 XX
 PS Claim 27; SEQ ID NO 636; 339pp + Sequence Listing; English.
 CC This invention describes a novel disease detection and treatment molecule
 CC polypeptide (MDPT) which has anti-inflammatory, immunosuppressive,
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephroprotective,
 CC antianemic, antipsoriatic and hepatoprotective activity. The polynucleotides
 CC and the polypeptides of the invention can be used for gene therapy.
 CC protein replacement therapy and are useful for treating a variety of
 CC diseases or conditions. These polypeptides or polynucleotides are
 CC particularly useful for diagnosing, treating or preventing cell
 CC proliferative disorders (e.g. cancers including adenocarcinoma,
 CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
 CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
 CC hepatitis. ABU11450-ABU11845 represent the MDPT polynucleotides encoded
 CC by ABU11450-ABU11845, described in the disclosure of the invention.
 CC NOTE: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 197 AA:
 Query Match 52.1%; Score 948; DB 24; Length 197;
 Best Local Similarity 98.9%; Pred. No. 3.7e-93;
 Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFSEDDPVVKKIALREIRMLKQLKH 60
 DB 6 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFSEDDPVVKKIALREIRMLKQLKH 65
 QY 61 PNLVNLIEVFRKRKHVLVFEYCDHTLLNELERNPGVADGVKSVLWQTLQALNFCGHIH 120
 DB 66 PNLVNLIEVFRKRKHVLVFEYCDHTLLNELERNPGVADGVKSVLWQTLQALNFCGHIH 125
 QY 121 NC:HRDIKPENILITKQGIKICDFGFAQILPGDAYTDYVATRYRAPELLVGTQYGS 180
 DB 126 NC:HRDIKPENILITKQGIKICDFGFAQILPGDAYTDYVATRYRAPELLVGTQYGS 185
 QY 181 SV 182
 DB 186 SV 187
 RESULT 13
 AAY90724
 ID AAY90724 standard; Protein: 566 AA.

XX AAY90724;
 AC 15-AUG-2000 (first entry)
 DT Rabbit KKIAME kinase SEQ ID NO:4.
 DE Rabbit; KKIAME kinase; learning-induced kinase; learning; memory;
 KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;
 KW identification.
 XX Oryctolagus cuniculus.
 XX WO2000020567-A2.
 XX 13-AER-2000.
 XX 01-OCT-1999; 99WO-US23010.
 XX 02-OCT-1998; 98US-0102906.
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX Thompson RF, Gomi H, Sun W;
 P WPI; 2000-328932/28.
 DR N-PSDB; AAA29745.
 DR
 XX Novel learning induced kinase polynucleotides and polypeptides, useful
 PT for the analysis of learning and memory, and for gene therapy -
 PT
 XX Claim 1; Fig 4; 64pp; English.
 CC The present sequence represents a learning-induced kinase, designated
 CC KKIAME kinase, which is isolated from rabbit brain tissue. KKIAME
 CC kinase is a cdc2-related kinase. The KKIAME kinase polynucleotides can
 CC be used to express recombinant protein for analysis, characterisation or
 CC therapeutic use, as markers for tissues in which the protein is
 CC preferentially expressed, as molecular weight markers on Southern gels,
 CC as chromosome markers or tags, to compare endogenous DNA sequences in
 CC patients to identify potential genetic disorders, as probes to hybridise
 CC and discover novel related sequences, as a source of PCR primers, and as
 CC an antigen to induce anti-DNA antibodies. The polypeptides can be used
 CC in assay to discover biological activity, to raise antibodies, as tissue
 CC markers and to isolate correlative receptors or ligands. The
 CC polynucleotides may also be used for gene therapy for the treatment of
 CC disorders which are mediated by KKIAME kinase.
 XX
 SQ Sequence 566 AA:
 Query Match 51.7%; Score 941.5; DB 21; Length 566;
 Best Local Similarity 49.7%; Pred. No. 9.7e-92;
 Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;
 QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFSEDDPVVKKIALREIRMLKQLKH 60
 DB 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFSEDDPVVKKIALREIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKRKHVLVFEYCDHTLLNELERNPGVADGVKSVLWQTLQALNFCGHIH 120
 DB 66 PNLVNLIEVFRKRKHVLVFEYCDHTLLNELERNPGVADGVKSVLWQTLQALNFCGHIH 120
 QY 121 NC:HRDIKPENILITKQGIKICDFGFAQILPGDAYTDYVATRYRAPELLVGTQYGS 179
 DB 121 NC:HRDIKPENILITKQGIKICDFGFAQILPGDAYTDYVATRYRAPELLVGTQYGS 180
 QY 180 SV182
 DB 186 SV 187
 RESULT 13
 AAY90724
 ID AAY90724 standard; Protein: 566 AA.

CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX
SQ Sequence 187 AA;

Query Match 43.5%; Score 791; DB 22; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.5e-76;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEKYEKLAKTGGSGYGVVFKCRNKTSQGVAVKKEVESEDDPVVKKIALREIRMLKQLKH 60
Db |||||
Qy 1 MEKYEKLAKTGGSGYGVVFKCRNKTSQGVAVKKEVESEDDPVVKKIALREIRMLKQLKH 60
Db |||||

Qy 61 PNLVNLIEVFRKRKRMHLVFEYCDHTLLNELEARNPNGVADGVKSVLWQTLQALNFCIH 123
Db |||||

Qy 61 PNLVNLIEVFRKRKRMHLVFEYCDHTLLNELEARNPNGVADGVKSVLWQTLQALNFCIH 120
Db |||||

Qy 121 NCIHRIKPENILITKQGIKICDFGFAQIL 151
Db |||||

Qy 121 NCIHRIKPENILITKQGIKICDFGFAQIL 151
Db |||||

Search completed: November 13, 2003, 14:18:47
Job time : 41.8369 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:17:22 ; Search time 15.7251 Seconds
(without alignments)
933.660 Million ce-1 updates/sec

Title: US-09-671-050-10

Perfect score: 1820

Sequence: 1 MEKYEKLAKTGGSGYGVVFK.....RKARNEGRRRRQCVLPKLS 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: C

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 40 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/iaa/5A.COMB.psp: *
2: /cgn2_6/prodata/1/iaa/5B.COMB.psp: *
3: /cgn2_6/prodata/1/iaa/6A.COMB.psp: *
4: /cgn2_6/prodata/1/iaa/6B.COMB.psp: *
5: /cgn2_6/prodata/1/iaa/PCUS.COMB.psp: *
6: /cgn2_6/prodata/1/iaa/backfiles1.psp: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1157.5	63.6	358	4	US-09-411-628-11
2	941.5	51.7	566	4	US-09-411-628-4
3	915.5	50.3	493	4	US-09-411-628-10
4	563	30.9	544	4	US-09-417-197-113
5	557	30.6	298	2	US-08-874-347-25
6	557	30.6	298	2	US-08-969-106-2
7	557	30.6	298	3	US-09-093-522-25
8	557	30.6	298	4	US-09-338-125-2
9	557	30.6	298	4	US-09-265-225D-14
10	557	30.6	544	4	US-09-417-197-115
11	556	30.5	298	4	US-09-411-628-115
12	554	30.4	297	1	US-08-176-620A-16
13	554	30.4	297	2	US-08-461-985-16
14	554	30.4	298	4	US-09-457-040B-29
15	551.5	30.3	294	2	US-08-874-347-26
16	551.5	30.3	294	3	US-09-093-522-26
17	547.5	30.1	270	2	US-07-857-224B-31
18	545	29.9	297	2	US-08-874-347-24
19	545	29.9	297	3	US-09-093-522-24
20	545	29.9	297	4	US-09-411-628-12
21	541	29.7	297	2	US-08-874-347-23
22	541	29.7	297	3	US-09-093-522-23
23	540	29.7	274	1	US-08-318-947A-20
24	540	29.7	274	2	US-08-795-303-20
25	538.5	29.6	300	2	US-08-874-347-10
26	538.5	29.6	300	3	US-09-093-522-10
27	526	28.9	292	1	US-08-154-915-2

28	526	28.9	292	2	US-08-464-517-38	Sequence 38, Appl
29	526	28.9	292	2	US-08-246-361A-38	Sequence 38, Appl
30	526	28.9	292	3	US-08-463-773-38	Sequence 38, Appl
31	526	28.9	292	5	PCT-US93-09945-2	Sequence 2, Appl
32	526	28.9	297	3	US-08-932-787B-21	Sequence 21, Appl
33	526	28.9	297	3	US-08-932-012C-21	Sequence 21, Appl
34	526	28.9	297	4	US-08-888-818C-21	Sequence 21, Appl
35	498.5	27.4	297	3	US-09-093-522-22	Sequence 22, Appl
36	498.5	27.4	297	3	US-08-874-347-22	Sequence 22, Appl
37	493.5	27.1	274	2	US-07-857-224B-30	Sequence 30, Appl
38	493.5	27.0	527	4	US-09-739-455-2	Sequence 2, Appl
39	489.5	26.9	282	1	US-08-318-947A-19	Sequence 19, Appl
40	489.5	26.9	282	2	US-08-795-303-19	Sequence 19, Appl
41	489	26.9	323	2	US-08-874-347-21	Sequence 21, Appl
42	489	26.9	323	3	US-09-093-522-21	Sequence 21, Appl
43	488.5	26.8	317	1	US-08-463-090B-9	Sequence 9, Appl
44	488.5	26.8	317	2	US-08-874-347-18	Sequence 18, Appl
45	488.5	26.8	317	3	US-09-093-522-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-411-628-11
; Sequence 11, Application US/09411528
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/03/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1999-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-628-11

Query Match	63.6%	Score 1157.5;	DB 4;	Length 358;
Best Local Similarity	64.7%	Pred. No. 9.5e-114;		
Matches 211:	Conservative 41;	Mismatches 41;	Indels 33;	Gaps 2;
Qy	1	MEKYEKLAKTGGSGYGVVFKCRNKTSGGVAVYKPFVSEDDPVVKKIALREIRMLKQLKH	60	
Db	2	MEKYEKIGIGESYGVVFKCRNRTGQVAKFKFLESEDDPVVKKIALREIRMLKQLKH	61	
Qy	61	PNLVNLIEVFRKRKGLVFEYCDHTLNLNELERNPNVGADVIVKSVLMTLOALNFCIH	120	
Db	62	PNLVNLIEVFRKRKGLVFEYCDHTLNLNELERNPNVGADVIVKSVLMTLOALNFCIH	121	
Qy	121	NCIHRDKPENILITKQIKICDGFQAQILI-PGDAYTDYVATWYRAPELLVGDQYGV	179	
Db	122	NCIHRDKPENILITKQIKICDGFQAQILI-PGDAYTDYVATWYRAPELLVGDQYGV	181	
Qy	180	SSVDIWAICVFAELLTGQPLWPGKSDVDOLYLIIRTVETGFRHVDQAGLELLSSDPP	239	
Db	182	PPVDWAICVFAELLTGQPLWPGKSDVDOLYLIIRTVETGFRHVDQAGLELLSSDPP	219	
Qy	240	AVASOSAGITGKLIIRHOSIFKSNQFPHGISIPEDEMTELEKFSVDHPVALNFMKGL	299	
Db	220	-----GDLIRHQVQVFTNQYFSGVKIPDPEDNEPLEKFPNISYPALGLKGL	269	
Qy	300	KNPDDRITCSQLLESYFDSFQEAQ	325	
Db	270	HMDPTERTCEQLLHPHPENIREIE	295	

RESULT 2

US-09-411-628-4
; Sequence 4, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411.628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102.906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-411-628-4

Query Match 51.7%; Score 941.5; DB 4; Length 566;
Best Local Similarity 49.7%; Pred. No. 1.2e-90;
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSCQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGEISYGVVFKCRNKTSCQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKKHMLVFEYCDHTLLNELEPNPVGADGVKSVLMQTLQALNFCCHI 120
DB 61 ENLVNLIEVCKKRWYLVFEFVDHTILDOLLEFPNGLDYQVVKYLFQINGIGFCHSH 120
QY 121 NC1HRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQYG 179
DB 121 NI1HRDIKPNILVQSQGWKLCDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQYG 179
QY 180 SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
DB 180 KAVDVAIGCLVTEMFNGEPLFGDSDDIDQYHMMCL----- 218
QY 240 AVASQAGITCKLIPRHQSIFKSGFPHGISIPEDMETLEKFSQVHPVALNFMKGCL 299
DB 240 AVASQAGITCKLIPRHQSIFKSGFPHGISIPEDMETLEKFSQVHPVALNFMKGCL 299
QY 300 KNPDDRLTCSQLLESSYF--DSF-----QEAQK-RKARN-----EGNRRRQ 341
DB 269 HVDPKRPFCALLHHDFFQMDGFAERFSQELQVKQDARNVSLSKSQNRKKEK 324

RESULT 3

US-09-411-628-10
; Sequence 10, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411.628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102.906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-628-10

Query Match 50.3%; Score 915.5; DB 4; Length 493;
Best Local Similarity 48.3%; Pred. No. 5.5e-86;

Matches 172; Conservative 68; Mismatches 69; Indels 47; Gaps 6;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSCQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGEISYGVVFKCRNKTSCQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKKHMLVFEYCDHTLLNELEPNPVGADGVKSVLMQTLQALNFCCHI 120
DB 61 ENLVNLIEVCKKRWYLVFEFVDHTILDOLLEFPNGLDYQVVKYLFQINGIGFCHSH 120
QY 121 NC1HRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQYG 179
DB 121 NI1HRDIKPNILVQSQGWKLCDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQYG 179
QY 180 SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
DB 180 KAVDVAIGCLVTEMFNGEPLFGDSDDIDQYHMMCL----- 218
QY 240 AVASQAGITCKLIPRHQSIFKSGFPHGISIPEDMETLEKFSQVHPVALNFMKGCL 299
DB 240 AVASQAGITCKLIPRHQSIFKSGFPHGISIPEDMETLEKFSQVHPVALNFMKGCL 299
QY 300 KNPDDRLTCSQLLESSYF--DSF-----QEAQK-RKARN-----EGNRRRQ 341
DB 269 HVDPKRPFCALLHHDFFQMDGFAERFSQELQVKQDARNVSLSKSQNRKKEK 324

RESULT 4

US-09-417-197-113
; Sequence 113, Application US/09417197
; Patent No. 6518221
; GENERAL INFORMATION:
; APPLICANT: Ole Thastrup, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417.197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 113
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CDK2-EGFP fusion
US-09-417-197-113

Query Match 30.9%; Score 563; DB 4; Length 544;
Best Local Similarity 37.2%; Pred. No. 1.1e-50;
Matches 124; Conservative 58; Mismatches 95; Indels 56; Gaps 8;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSCQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MENFQVKEIGEGTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRREISLLKELNH 60
QY 61 PNLVNLIEVFRKKHMLVFEYCDHTLLNELEPNP-NGVADGVKSVLMQTLQALNFCCHI 119
DB 61 PNLVNLIDVHTENKLYNVFEFLHQDLKKFMDASALTGIP-PLIKSVLFOLLOGLAFCHS 120
QY 120 HNC1HRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQY 176
DB 121 HRVLRDLKPNQNLINTEGAIKLADFLARAFGVPRVRYTHEVVTWLYRAPEILLGSKY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTL----- 217
DB 180 STAVDIWSLGCIFABMYTRRALFGDSEIDQLFRIFRLTGTPTDEVPVWPGVTSMPDYKPSF 240
QY 218 --VETGFRHV---DOAGLELJTSS---DPPAVASQAGITCKLIPRHQSIFKSGNFFH 267
DB 240 PKWAKQDFSKVVPPLDEGRLSLQMLHYDPNKRISAKAAL-----AHPFFQ 287
QY 268 GISIPEPE-----DMETL-----EKFSDVHPV 290

DB 288 DVTKPVPHRLWDPVPVATWWSKGBELFTGVWP: 323

RESULT 5
US-08-874-347-25
; Sequence 25, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leaf, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CD2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; TITLE OF INVENTION: CARINIC
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version: 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-874-347-25

Query Match 30.6%; Score 557; DB 2; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
QY 1 MEKYEKLAKTSGSGVGVFKCRNKTSGOVAVKVFVSEDDPVYKKIALREIRMLKOLKH 60
DB 1 MENFOKEKIGEGTYGVVYKARNKLTGEWALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRGMHLEVEECCHTLLNELRNP-NGVADGVKSVLWOTLQALNECHI 119
DB 61 PNLVNLIDVHTENKLYLVFELHODLKKFMDASALTGIPPLIKSYLFLQLGLAFCHS 120
QY 120 HNCIHRDKPENILITKQIICIDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQY 178
DB 121 HRVLRDLKPNQLLINTEGAIKLADFGLABAGFVPVRYTHVVTWLYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLMPKSKSDVDQLYLIIRTL----- 217
DB 181 STAVDIMSGLGCIFAEMVTRRALFPDSEIDQLFRIFRLTGPDEWVPGVTSMPDYKPSF 240
QY 218 ---VETGRHV-----DQAGLELLTSS---DPPAVASQAGITGKLIHRHOSIFKSNQFPF 267
DB 241 PKWARQDFSKVVPPLDEGRSLLSQMLHYDPNKRISAKAAL-----AHPFFQ 287

QY 268 GISIDEP 274
DB 288 DVTKPVP 294
RESULT 6
US-09-969-106-2
; Sequence 2, Application US/08969106
; Patent No. 5986055
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-969-106-2

Query Match 30.6%; Score 557; DB 2; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
QY 1 MEKYEKLAKTSGSGVGVFKCRNKTSGOVAVKVFVSEDDPVYKKIALREIRMLKOLKH 60
DB 1 MENFOKEKIGEGTYGVVYKARNKLTGEWALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRGMHLEVEECCHTLLNELRNP-NGVADGVKSVLWOTLQALNECHI 119
DB 61 PNLVNLIDVHTENKLYLVFELHODLKKFMDASALTGIPPLIKSYLFLQLGLAFCHS 120
QY 120 HNCIHRDKPENILITKQIICIDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQY 178
DB 121 HRVLRDLKPNQLLINTEGAIKLADFGLABAGFVPVRYTHVVTWLYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLMPKSKSDVDQLYLIIRTL----- 217
DB 181 STAVDIMSGLGCIFAEMVTRRALFPDSEIDQLFRIFRLTGPDEWVPGVTSMPDYKPSF 240
QY 218 ---VETGRHV-----DQAGLELLTSS---DPPAVASQAGITGKLIHRHOSIFKSNQFPF 267
DB 241 PKWARQDFSKVVPPLDEGRSLLSQMLHYDPNKRISAKAAL-----AHPFFQ 287
QY 268 GISIDEP 274

```
Db      288 DVTKPVP 294      ::|||
::|||
241 PKWARQDFSKVPPPLDGDGRSLLSQMLHYDPNKRISAKAAL-----AHPFQ 287
QY      268 GISPEP 274      ::|||
::|||
Db      288 DVTKPVP 294      ::|||
::|||

RESULT 8
US-09-338-125-2
; Sequence 2, Application US/09338125
; Patent No. 6521412
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Mandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,125
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9695
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-093-522-25

Query Match      30.6%; Score 557; DB 3; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY      1 MEKYEKLAKTGECSGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIMLKQLKH 6C
Db      2 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY      61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 119
Db      61 PNLVLLDVHTENKLYLVFEFLHODLKKFMDASALTGIPPLIKSVLFOLLQGLAFCHS 120
QY      120 HNCIHRDIKPNILITKGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGSQTQY 178
Db      121 HRVLHRLDKPQNLLINTEGAIKKADFGLARAFGVPRVITYHEVVTLYWYRAPEILLGSKYY 180
QY      179 GSSVDIWAIGCVFAELLTGQCPWPKSDVQOLYLIIRTL-----217
Db      181 STAVDIWSLGCIFAEVMTRRALFPDSEIDQLFRITLGTPEVWVPGVTSMPDYKPSF 240
QY      218 ---VETGRHV-----DQAGLELLTSS---OPPAVASQASAGITGKLIIRHQSIFKSKGFPH 267
```

```
Db      241 PKWARQDFSKVPPPLDGDGRSLLSQMLHYDPNKRISAKAAL-----AHPFQ 287
QY      268 GISPEP 274      ::|||
::|||
Db      288 DVTKPVP 294      ::|||
::|||

RESULT 8
US-09-338-125-2
; Sequence 2, Application US/09338125
; Patent No. 6521412
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Mandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,125
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-338-125-2

Query Match      30.6%; Score 557; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY      1 MEKYEKLAKTGECSGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIMLKQLKH 60
Db      1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY      61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 119
Db      61 PNLVLLDVHTENKLYLVFEFLHODLKKFMDASALTGIPPLIKSVLFOLLQGLAFCHS 120
QY      120 HNCIHRDIKPNILITKGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGSQTQY 178
Db      121 HRVLHRLDKPQNLLINTEGAIKKADFGLARAFGVPRVITYHEVVTLYWYRAPEILLGSKYY 180
QY      179 GSSVDIWAIGCVFAELLTGQCPWPKSDVQOLYLIIRTL-----217
Db      181 STAVDIWSLGCIFAEVMTRRALFPDSEIDQLFRITLGTPEVWVPGVTSMPDYKPSF 240
```

```
QY 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
Db 241 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 287
QY 268 GISIEPEP 274
Db 288 DVTKPVP 294

RESULT 9
US-09-266-225D-14
; Sequence 14, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; INTERACTING PROTEINS
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-14

Query Match 30.6%; Score 557; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKXKFESEDDPVVVKXIALREIRMLKQLKH 60
Db 1 MENFOKVEKIGEGTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
QY 307 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 366
Db 307 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 366
QY 120 HNCIHRDIKPNILITKQIICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGTOY 178
Db 120 HNCIHRDIKPNILITKQIICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGTOY 178
QY 367 HRVLRDLKPNQLLNTGAIKLDGFLARAFGVPTVTHEVVTLYRAPEILGSKYY 426
Db 367 HRVLRDLKPNQLLNTGAIKLDGFLARAFGVPTVTHEVVTLYRAPEILGSKYY 426
QY 179 GSSVDIWAICGVFAELLTGQPLWPKGSDVDQYLIIRTL----- 217
Db 179 GSSVDIWAICGVFAELLTGQPLWPKGSDVDQYLIIRTL----- 217
QY 427 STAVDIMSGLGCFABWVTRRALFPDSEIDQLFRIFRITLGTPEVVPVGTSMFYKPSF 486
Db 427 STAVDIMSGLGCFABWVTRRALFPDSEIDQLFRIFRITLGTPEVVPVGTSMFYKPSF 486
QY 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
Db 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
QY 487 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 533
Db 487 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 533
QY 268 GISIEPEP 274
Db 534 DVTKPVP 540

RESULT 11
US-09-411-628-13
; Sequence 13, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(298)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-411-628-13

Query Match 30.5%; Score 556; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 2.4e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKXKFESEDDPVVVKXIALREIRMLKQLKH 60
Db 1 MENFOKVEKIGEGTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
```

```
QY 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
Db 241 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 287
QY 268 GISIEPEP 274
Db 288 DVTKPVP 294

RESULT 9
US-09-266-225D-14
; Sequence 14, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; INTERACTING PROTEINS
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-14

Query Match 30.6%; Score 557; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKXKFESEDDPVVVKXIALREIRMLKQLKH 60
Db 1 MENFOKVEKIGEGTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
QY 307 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 366
Db 307 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 366
QY 120 HNCIHRDIKPNILITKQIICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGTOY 178
Db 120 HNCIHRDIKPNILITKQIICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGTOY 178
QY 367 HRVLRDLKPNQLLNTGAIKLDGFLARAFGVPTVTHEVVTLYRAPEILGSKYY 426
Db 367 HRVLRDLKPNQLLNTGAIKLDGFLARAFGVPTVTHEVVTLYRAPEILGSKYY 426
QY 179 GSSVDIWAICGVFAELLTGQPLWPKGSDVDQYLIIRTL----- 217
Db 179 GSSVDIWAICGVFAELLTGQPLWPKGSDVDQYLIIRTL----- 217
QY 427 STAVDIMSGLGCFABWVTRRALFPDSEIDQLFRIFRITLGTPEVVPVGTSMFYKPSF 486
Db 427 STAVDIMSGLGCFABWVTRRALFPDSEIDQLFRIFRITLGTPEVVPVGTSMFYKPSF 486
QY 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
Db 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
QY 487 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 533
Db 487 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 533
QY 268 GISIEPEP 274
Db 534 DVTKPVP 540

RESULT 11
US-09-411-628-13
; Sequence 13, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(298)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-411-628-13

Query Match 30.5%; Score 556; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 2.4e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKXKFESEDDPVVVKXIALREIRMLKQLKH 60
Db 1 MENFOKVEKIGEGTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
```

```
QY 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
Db 241 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 287
QY 268 GISIEPEP 274
Db 288 DVTKPVP 294

RESULT 9
US-09-266-225D-14
; Sequence 14, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; INTERACTING PROTEINS
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-14

Query Match 30.6%; Score 557; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKXKFESEDDPVVVKXIALREIRMLKQLKH 60
Db 1 MENFOKVEKIGEGTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
QY 307 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 366
Db 307 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 366
QY 120 HNCIHRDIKPNILITKQIICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGTOY 178
Db 120 HNCIHRDIKPNILITKQIICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGTOY 178
QY 367 HRVLRDLKPNQLLNTGAIKLDGFLARAFGVPTVTHEVVTLYRAPEILGSKYY 426
Db 367 HRVLRDLKPNQLLNTGAIKLDGFLARAFGVPTVTHEVVTLYRAPEILGSKYY 426
QY 179 GSSVDIWAICGVFAELLTGQPLWPKGSDVDQYLIIRTL----- 217
Db 179 GSSVDIWAICGVFAELLTGQPLWPKGSDVDQYLIIRTL----- 217
QY 427 STAVDIMSGLGCFABWVTRRALFPDSEIDQLFRIFRITLGTPEVVPVGTSMFYKPSF 486
Db 427 STAVDIMSGLGCFABWVTRRALFPDSEIDQLFRIFRITLGTPEVVPVGTSMFYKPSF 486
QY 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
Db 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQSGITGKLI PRHQSIFKSNQFFH 267
QY 487 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 533
Db 487 PKWARQDSKVVPPLDEDGRSLQYLHCPNKRISAKAAL-----AHPEFFQ 533
QY 268 GISIEPEP 274
Db 534 DVTKPVP 540

RESULT 11
US-09-411-628-13
; Sequence 13, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(298)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-411-628-13

Query Match 30.5%; Score 556; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 2.4e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKXKFESEDDPVVVKXIALREIRMLKQLKH 60
Db 1 MENFOKVEKIGEGTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVIKSVLWQTLQALNECHI 119
```

DB 61 PNIVKLELVHTENKJLVFEFLHQLDKKFDASALTGIPLIKSYLFQLQLGAFCHS 120
QY 120 HNCIHRDIKENILITKQGIKICDFGFAQIL-IPGDAYTQYVATRWYRAPELLVGDYQ 178
DB 121 HRVLRHLKPNLINTEGAIAKADFGARAFGVFEVITYTHEVVTLWYRAPEILLGCKY 180
QY 179 GSSVDIWAIGCVFAELLITGQPLMPGKSDVQQLYLIIRTL----- 217
DB 181 STAVDINSLGCIFAEMVTTRRALPGDSEIQDLFRITLGTPEVWPGVTSMPYKPSF 240
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASOSAGITGKLIPIRHOS*FKSNGGFH 267
DB 241 PKWARQDSKVVPBLDBGRSLLSQMLHYDHNKRISAKAAJ-----AHPFFQ 287
QY 268 GISIEP 274
DB 286 DVTXFPV 294

RESULT 12
US-08-176-620A-16
; Sequence 16, Application US/08176620A
; Patent No. 5595904

GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176.620A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-176-620A-16

Query Match 30.4%; Score 554; DB 1; Length 297;
Best Local Similarity 40.5%; Pred. No. 3.8e-50;
Matches 118; Conservative 58; Mismatches 73; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGEISGVGVVFKCRNKTSGQVAVKVF-VESEDDPVVKKIALREIRMLKQLK 59
DB 1 MEDYTKIEKIGEGYGVVYGRHKHTTQGVVAMKKIRLESEEG-VPSTAIRISJLKEUR 59
QY 60 HPNLVNI--EVFRKRKMKHLVFEYCDHTLLNELEARNPNG--VADGV-KSVJWQTLQALNFC 117

DB 60 HPNIVSLQDVLWQDSRLYLIFEFSMDLKKYLSIPPGQYMDSSLVKSXYLQIGIVFC 119
QY 118 HHNCIHRDIKENILITKQGIKICDFGFAQIL-IPGDAYTQYVATRWYRAPELLVGDY 176
DB 120 HSRVLRHLKPNLINTEGAIAKADFGARAFGVFEVITYTHEVVTLWYRAPEILLGSA 179
QY 177 QYSSVDIWAIGCVFAELLITGQPLMPGKSDVQQLYLIIRTL----- 217
DB 180 RYSPVDIWSIGTIFAELATKKPLPHGDSIDQDLFRIFRALGTPNNEVNEVESLQDYKN 239
QY 218 -----VETGFRHVQDQAGLELLTSS---DPPAVASOSAGITGKLIPIRH 256
DB 240 TFPKMKPGSLASHVKNLDEGLDLSKMLIYDP-----AKRISGMALNH 284

RESULT 13
US-08-461-985-16
; Sequence 16, Application US/08461985
; Patent No. 5872006

GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461.985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176.620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-461-985-16

Query Match 30.4%; Score 554; DB 2; Length 297;
Best Local Similarity 40.5%; Pred. No. 3.8e-50;
Matches 118; Conservative 58; Mismatches 73; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGEISGVGVVFKCRNKTSGQVAVKVF-VESEDDPVVKKIALREIRMLKQLK 59
DB 1 MEDYTKIEKIGEGYGVVYGRHKHTTQGVVAMKKIRLESEEG-VPSTAIRISJLKEUR 59
QY 60 HPNLVNI--EVFRKRKMKHLVFEYCDHTLLNELEARNPNG--VADGV-KSVJWQTLQALNFC 117

```

Db 60 HPIVSLQDVXQDSRLYLIFELSLMDLKKYLDSPGQVMDSSLVKSYIYQIQGVFC 129
Qy 118 HINCHIRDIKPENILITKQGIKICDFGAQIL-IPGDAYTVATRWYRAPELLVGET 176
Db 120 HSRVYLHRLKPNQLLIDDKGTIKADFLARAFGIPRVYTHVYVTLWYASPEVLGSA 179
Qy 177 QYSSVDIWAIGCVFAELLTGQPLWPKGSVDQVLYLIRTL----- 217
Db 180 RYSTPVDIWSIGTIFAEALATKPKPLFGHGDSEIDQLFRIFALGTGPNNEVWPEVESLDQYKN 239
Qy 218 -----VETGPRHVDAQ:ELLTSS---DPPAVASQAGITCKLIPRH 256
Db 240 TFPKWKPSLASHVKNLQDGLDLSKMLIYDP-----AKRISGKVALNH 284

RESULT 14
US-09-457-040B-29
; Sequence 29, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 298
; TYPE: P3T
; ORGANISM: Human
US-09-457-040B-29

```

```

Query Match 30.4%; Score 554; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 3.8e-50;
Matches 118; Conservative 54; Mismatches 85; Indels 46; Gaps 6;

Qy 1 MEKYELAKTGGSGYGVFKCRNKTSGQVAVKFESEDDPVVKKIALREIRMLKQKH 60
Db 1 MENFOKVEKIGESTYGVVYKARNKTGGEVALKKIRLDITETGVFSTAIRISLKLNLH 60
Qy 61 PNLVNLIEVFRKRKRMHVFVEYCDHTLNLERNP-NGVADGVKSVLWOTLOALNFC 119
Db 61 PNLVNLVITENKLYLVFELHDLKKFMDASALTGIPFLIKSLFQLLOGLAFCHS 120
Qy 120 HNCIHRDIKPENILITKQGIKICDFGAQIL-IPGDAYTVATRWYRAPELLVGETQY 178
Db 121 HRVLHRLKPNQLLINTEGAIKADFLARAFGVEVRYTHVYVTLWYRAPELLVGCYV 180
Qy 179 GSSVDIWAIGCVFAELLTGQPLWPKGSVDQVLYLIRTL----- 217
Db 181 STAVDIWSGCIFAEMVTRRALFGDSEIDQGFIRFLGTGPDVEWPGVTSNPDYKSP 240
Qy 218 ---VETGPRHV-----DQAGLELLTSS---DPPAVASQAGITCKLIPRHOSIFKSN 267
Db 241 PKWARQDSKVVPVLDGDSRLSQVLYHDPNKRISAKAAL-----AHFFPQ 267
Qy 268 GISIPEP 274
Db 288 DVTKPVP 294

```

```

RESULT 25
US-08-874-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leaf, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.

```

```

; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/C55001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-347-26

Query Match 30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 34.8%; Pred. No. 6.9e-50;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;

Qy 1 MEKYELAKTGGSGYGVFKCRNKTSGQVAVKFESEDDPVVKKIALREIRMLKQKH 60
Db 1 MEQYEKEEKIGESTYGVVYRAEDKVTNETIALKKIRLQEQDEGVSTAIRISLKKEMHH 60
Qy 61 PNLVNLIEVFRKRKRMHVFVEYCDHTLNL-----ELERNPGVADGVKSVLWOTLOAL 114
Db 61 GNIVLHDVHSEKRIYLVFEYLDLKKFMDSCPEFAKNPT-----LKSILYQILRGV 115
Qy 115 NFCHINCHIRDIKPENILITKQ-GIIKICDFGAQIL-IPGDAYTVATRWYRAPELL 172
Db 116 AXCHSHRVLHRLKPNQLLIDERTNALKLADFLARAFGIPVTFTHVYVTLWYRAPEIL 175
Qy 173 VQDTQYSSVDIWAIGCVFAELLTGQPLWPKGSVDQVLYLIRTLVETGFRHVDQAGLEL 232
Db 176 LGSROYSTPVDVMSYGCIFAEWVKNOKPLFGDSEIDELFKIFRVLGTP----- 223
Qy 233 LSSSDPPAVASQAGITCKLIPRHOSIFKSNPFHFGISIPEDMETLEEFSDVHPVAL 292
Db 224 -NEQSWPGVSS-----LPDYKSAFPK-----WQQLDATTI---VPTLDPAGL 261
Qy 293 NFMKGCLXNPDORLTCSQLLESYFDSFQBAQ 325
Db 262 DLLSKMLAYEPNKRITARQALEHEVFKDLEMVQ 294

```

```

Search completed: November 13, 2003, 14:21:59
Job time : 16.7251 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: November 13, 2003, 14:16:37 ; Search time 14.6767 Seconds
(without alignment)
2273.702 Million cell updates/sec

Title: US-09-671-050-10

Perfect score: 1820
Sequence: 1 MEKYELAKTGGSGYGVFK.....PKARKEGRNRPPQVLPLKS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*

- 1: PIR1:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1157.5	63.6	358	S23383	protein kinase (EC
2	1149.5	63.2	376	S22745	serine/threonine p
3	565	31.0	305	S23382	protein kinase (EC
4	563	30.9	302	I50474	protein kinase (EC
5	558.5	30.7	292	S40021	protein kinase (EC
6	555	30.5	297	A37571	protein kinase (EC
7	554	30.4	298	A41227	protein kinase (EC
8	552.5	30.4	294	B40444	protein kinase (EC
9	551.5	30.3	294	S21095	protein kinase (EC
10	551.5	30.3	294	S22440	protein kinase (EC
11	546.5	30.0	294	A40444	protein kinase (EC
12	546.5	30.0	294	I49271	CELL DIVISION CON
13	545	29.9	297	A29539	protein kinase (EC
14	545	29.9	298	A44378	protein kinase (EC
15	542	29.8	297	I45977	cyclin-dependent k
16	541	29.7	297	S24913	protein kinase (EC
17	540.5	29.7	297	S12009	protein kinase cdc
18	540.5	29.7	302	B44349	protein kinase (EC
19	540	29.7	302	A44349	protein kinase (EC
20	539	29.6	303	S06011	protein kinase (EC
21	537	29.5	288	S42566	protein kinase (EC
22	536.5	29.5	311	S36619	protein kinase (EC
23	535.5	29.4	294	JQ2243	protein kinase (EC
24	535.5	29.4	294	S57928	protein kinase (EC
25	535.5	29.4	294	S42049	protein kinase (EC
26	532	29.2	301	S19209	protein kinase (EC
27	531	29.2	292	A46365	protein kinase (EC
28	530.5	29.1	294	S31332	protein kinase (EC
29	529	29.1	292	I49592	protein kinase (EC

30	529	29.1	292	1	A45091	protein kinase (EC
31	527.5	29.0	308	1	S53538	protein kinase (EC
32	527	29.0	297	1	A36074	protein kinase (EC
33	527	29.0	301	1	S42101	protein kinase (EC
34	526.5	28.9	291	2	S23386	protein kinase (EC
35	526	28.9	293	2	JE0374	cyclin-dependent k
36	526	28.9	294	2	S51008	protein kinase (EC
37	526	28.9	296	2	S24386	protein kinase PHO
38	524	28.8	302	1	OKBY85	protein kinase cdc
39	521.5	28.7	302	2	T17115	CDK-activating pro
40	519.5	28.5	346	2	A54820	protein kinase (EC
41	518	28.5	292	2	S22441	MO15/CDK-activatin
42	517.5	28.4	346	2	A56231	protein kinase (EC
43	516.5	28.4	314	1	S12007	protein kinase (EC
44	514	28.2	293	2	T02922	protein kinase (EC
45	514	28.2	346	1	I78840	protein kinase (EC

ALIGNMENTS

RESULT 1

S23383
protein kinase (EC 2.7.1.37) cdc2-related KXIALRE - human
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999
C:Accession: S23383; S22744
R:Myerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; T
EMBO J. 11, 2909-2917, 1992
A:Title: A family of human cdc2-related protein kinases.
A:Reference number: S23382; MUID:92347325; PMID:1639063
A:Accession: S23383
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-358 <MEY>
A:Cross-references: EMBL:X66358; NID:G36614; PIDN:CAA47002.1; PID:G36615
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:3-278/Domain: protein kinase homology <KIN>
F:11-19/Region: protein kinase ATP-binding motif
F:34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match	63.6%	Score 1157.5;	DB 2;	Length 358;
Best Local Similarity	64.7%	Pred No. 3.1e-46;		
Matches	211;	Conservative	41;	Mismatches 41; Indels 33; Gaps 2;
QY	1	MEKYELAKTGGSGYGVFKCRNKTSGQWAVKFESEDDPVVKKIALREIRMLKQLKH	60	
Db	2	MEKYELKIGEGSGYGVFKCRNRTGQIVAKKFLSEDDPVVKKIALREIRMLKQLKH	61	
QY	61	PNLVNLIIEVFRKRKHVPEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH	120	
Db	62	PNLVNLIIEVFRKRKHVPEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH	121	
QY	121	NCIHRDIPENILITKQIICKIDFGFAQIIL-PODAYTDYVATRWYNAPELLVGDYQYG	179	
Db	122	NCIHRDIPENILITKQIICKIDFGFAQIIL-PODAYTDYVATRWYNAPELLVGDYQYG	181	
QY	180	SSVDVNIICCVFAE::TCQPLWPGSDVDVLYLITLVE-GFRHVDQAGLELLTSSDPP	239	
Db	192	FPVDVNIICCVFAE::TCQPLWPGSDVDVLYLITLVE-GFRHVDQAGLELLTSSDPP	219	
QY	240	AVASQAGITGLIPRHQSIFKSNGFPHGISIPEDMETLEEFSDVHPVAFNMKGCL	299	
Db	220	-----GDILIPHQVQFSTNQYFGVXIPDPEDMEPLELAFPNISYPALGLLKGCL	269	
QY	300	KMPDRLTCSQLLESSYFDSFQEAQ	325	
Db	270	HMDPTERTLTCQLLHPYPENIREIE	295	

RESULT 2

S22745

serine/threonine protein kinase KKIAURE (EC 2.7.1.1) - human

C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Dec-1997
C:Accession: S22745

R:Mayerson, M.L. Data Library, May 1992

Submitted to the EMBL Data Library, May 1992

A:Reference number: S22743

A:Accession: S22745

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-376 <MEY>

A:Cross-references: EMBL:X66359

C:Genetics:

A:Introns: 152/3; 170/3

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:3-296/Domain: protein kinase homology <KIN>

F:11-19/Region: protein kinase ATP-binding motif

Query Match 63.2%; Score 1149.5; DB 2; Length 376;
Best Local Similarity 61.3%; Pred. No. 7.5e-46;
Matches 211; Conservative 41; Mismatches 41; Indels 51; Gaps 2;

QY 1 MEKYEKLAKTGEISGVVFKCRNKTSGOVAVKVESEDDPVYKXIALRE-RMLKQLXH 60

DB 2 MEKYEKIGKIGESGVVFKCRNRTDQIVAIKFKLESEDDPVYKXIALRE-RMLKQLXH 61

QY 61 PNLVNLIEVFRKRKHVLVEYCDHTLLNELERNPNNG-VADGVIKSVLWQTLQALNLF 120

DB 62 PNLVNLIEVFRKRKHVLVEYCDHTVJHEDRYQGVPEHLVKSITWQTLQAVNFCRH 121

QY 121 NCIHRLDKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRYRAPELLVGD 162

DB 122 NCIHRLDKPENILITKHSVILKICDFGFAQLFQPSAAVCFPCS-TGXITGPDYTDYV 182

QY 162 ATRWYRAPELLVGDITQYSSVDVWAGCVFAELLTGQPLKPGKSDVDQVLIIRTLVETG 221

DB 182 ATRWYRAPELLVGDITQYGPVDVWAGCVFAELLTGQPLKPGKSDVDQVLIIRTLV 237

QY 222 FRHVQAGLELTSSDDPPAVASQASGITKLIIPRHQSIFKSNFFHGHS-PEPEDMETLE 281

DB 238 -----GDLIPRHQVFTSTNOYFSGVKIPQPEDMEPLE 269

QY 282 EKESDVHPVALNPMKGLKKNPDRLTCSQLLESYFDSFCEAQ 325

DB 270 LKPNYSYPAZGLKGLCHMDPTERTCEGLJHHPVFENIRE-E 313

RESULT 3

S23382

Protein kinase (EC 2.7.1.37) cdk3 - human

C:Species: Homo sapiens (man)

C:Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 18-Jun-1999

C:Accession: S23382; S22743

R:Mayerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Harlow, E.; Tsai

EMBO J 11, 2909-2917, 1992

A:Title: A family of human cdc2-related protein kinases.

A:Reference number: S23382; MUID:92347325; PMID:1639063

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-305 <MEY>

A:Cross-references: EMBL:X66357; NID:g36612; PID:CAA47001.1; PID:g36613

C:Genetics:

A:Gene: CDK3

A:Cross-references: GDB:283456

A:Map position: 12q13-12q13

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase

F:2-255/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase ATP-binding motif

F:33,51,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 31.0%; Score 565; DB 2; Length 305;
Best Local Similarity 36.8%; Pred. No. 2.1e-19;
Matches 124; Conservative 66; Mismatches 99; Indels 48; Gaps 9;

QY 1 MEKYEKLAKTGEISGVVFKCRNKTSGOVAVKVF---VESEDDPVYKXIALREIRMLKQ 57

DB 1 MDMFKVEKIGEGTYGVYKAKNRETQGLVAKIRLDLEMEGVF---STAIREISLLKE 57

QY 58 LKHPNLVNLIEVFRKRKHVLVEYCDHTLLNELERNPNNG-VADGVIKSVLWQTLQALNLF 116

DB 58 LKHPNLVNLIEVFRKRKHVLVEYCDHTLLNELERNPNNG-VADGVIKSVLWQTLQALNLF 117

QY 117 CHHNCHIRDKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRYRAPELLVGD 175

DB 118 CHSHRVHRLDKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRYRAPELLVGD 177

QY 176 TQYSSVDVWAGCVFAELLTGQPLKPGKSDVDQVLIIRTLVETGFRHVDQAGLELLTS 235

DB 178 KFYTTAVDIMSIGCIFAEVMVTRKALFPDSEIDQIFRML----- 219

QY 236 SDPPAVASQSA--GITGLIPRHQSIFKSNFFHGHSIPEPE-DMETLEKFSVDVHPVAL 292

DB 223 ---GTPSEDTPGV---QPDYKGSF-----PKWTRKGLSEIVPNLEPEGR 260

QY 293 NPMKGLKKNPDRLTCSQLLESYFDSFCEAQIKRK 329

DB 261 DLEMCLLQYDPSQRTAKTALAHPYFSSPEPSPAARQ 297

RESULT 4

IS0474

Protein kinase (EC 2.7.1.37) cdc2 [similarity] - goldfish

C:Species: Carassius auratus (goldfish)

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000

C:Accession: IS0474

R:Kajitara, H.; Yanashita, M.; Katsu, Y.; Nagahama, Y.

Dev. Growth Differ. 35, 647-654, 1993

A:Title: Isolation and characterization of goldfish cdc2, a catalytic component of matu

A:Reference number: IS0474

A:Accession: IS0474

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-302 <KAC>

A:Cross-references: GB:D17758; NID:g471097; PID:BAA04605.1; PID:g471098

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase

F:2-256/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase ATP-binding motif

F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.9%; Score 563; DB 1; Length 302;

Best Local Similarity 43.2%; Pred. No. 2.6e-19;

Matches 117; Conservative 60; Mismatches 71; Indels 36; Gaps 6;

QY 1 MEKYEKLAKTGEISGVVFKCRNKTSGOVAVKVF---VESEDDPVYKXIALREIRMLKQ 59

DB 1 MDDYLKIEKIGEGTYGVYKGRNKTGQWAMKIRLESEEG-VPTAVREISLLKELQ 59

QY 60 HPNLVNLIEVFRKRKHVLVEYCDHTLLNELERNPNNG---VADGVIKSVLWQTLQALNFC 117

DB 60 HPNLVNLIEVFRKRKHVLVEYCDHTLLNELERNPNNG---VADGVIKSVLWQTLQALNFC 119

QY 119 HINCHIRDKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRYRAPELLVGD 176

DB 120 HCRRLVHRLDKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRYRAPELLVGD 179

QY 177 QYSSVDVWAGCVFAELLTGQPLKPGKSDVDQVLIIRTL----- 217

DB 180 RYSTPVDVWISGTTFAELATKQIFHGDSEIDQIFRITLGTPTNNNEWPDVDSLSLDYKN 239

QY 218 -----VETGFRHVDQAGLELLTS---DPPAVASQASGIT 249

DB 240 TFPKWSGNLASTVKNLDKNGIDLLTKWLIYDPPKRISARQAMT 283

RESULT 5

S40021
Protein kinase (EC 2.7.1.37) cdc2 homolog - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
A:Accession: S40021, S39071
R:Michaelis, C.E.; Weeks, G.
submitted to the EMBL Data Library, August 1992
A:Description: The unicellular organism Dictyostelium discoideum possesses a highly related
A:Reference number: S40021
A:Accession: S40021
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-292 <MI2>
A:Cross-references: EMBL:L00652; NID:g167695; PIDN:AAA16056.1; PID:g167695
R:Michaelis, C.; Weeks, G.
Biochim. Biophys. Acta 1179, 117-124, 1993
A:Title: The isolation from a unicellular organism, Dictyostelium discoideum, of a highly
A:Reference number: S39071; MUID:94032415; PMID:8218353
A:Accession: S39071
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-18, 'Y', '20-188, 'G', '190-292 <MI2>
A:Cross-references: EMBL:L00852
C:Complex: in various organisms, cdc2 has been identified as a component of the M-phase
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine
F:2-254/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,126,128/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.7%; Score 558.5; DB 2; Length 292;
Best Local Similarity 35.0%; Pred. No. 4e-19;
Matches 115; Conservative 66; Mismatches 101; Indels 47; Gaps 6;

QY 1 MEKYELKATGEGSGVGVFKCNKTSQGVAVKFESEDDPVVKKIALREIMLKQLKH 60
DB 1 MEKYEKELKATGEGSGVGVFKCNKTSQGVAVKFESEDDPVVKKIALREIMLKQLKH 60
QY 61 PNLVNLIEFRKRKMHVLFVEYCDHTLLNELE-RNPNGVADGVKSVLWQTLQALNFCIH 120
DB 61 PNLVNLIEFRKRKMHVLFVEYCDHTLLNELE-RNPNGVADGVKSVLWQTLQALNFCIH 120
QY 121 NCIHRIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDVYATRWYRAPELLVGDQYQ 179
DB 121 NCIHRIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDVYATRWYRAPELLVGDQYQ 179
QY 121 RVLHRDLKPNQLLNSDCAIKLADFGARAFGVPVRYTSHEVVTLYWYRAPELLVGDQYQ 180
DB 121 RVLHRDLKPNQLLNSDCAIKLADFGARAFGVPVRYTSHEVVTLYWYRAPELLVGDQYQ 180
QY 180 SSVDVMAIGCVFAELLTGQPLWPKSDVDQVLYLIRTL-VEVTFGRHVDQAGLELLTSDDP 239
DB 180 SSVDVMAIGCVFAELLTGQPLWPKSDVDQVLYLIRTL-VEVTFGRHVDQAGLELLTSDDP 239
QY 181 TPIDVWSALCIFAEMASGRPLFGSGTSDQVFRIFKILGTP-----NEESWP 227
DB 181 TPIDVWSALCIFAEMASGRPLFGSGTSDQVFRIFKILGTP-----NEESWP 227
QY 240 AVASOSAGITGKLIPRHQSIF-----XSGFFHC-SIPEPEDMETLEEKFSDEHPVALN 293
DB 240 AVASOSAGITGKLIPRHQSIF-----XSGFFHC-SIPEPEDMETLEEKFSDEHPVALN 293
QY 228 SITE-----LPEYKTFDPVHAHQUS-VHG-----LDEK-----GLN 260
DB 228 SITE-----LPEYKTFDPVHAHQUS-VHG-----LDEK-----GLN 260
QY 294 FMKGLKKNPDCDLTCSQLLESSYEDSFQ 322
DB 294 FMKGLKKNPDCDLTCSQLLESSYEDSFQ 322

RESULT 6

A37871
Protein kinase (EC 2.7.1.37) cdk2 - African clawed frog
N:Alternate names: cell division control protein CDC2 homo.log Eg:
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 18-Jun-1999
A:Accession: A37871; S15866; I51662; S14410
R:Paris, J.; Le Guellec, R.; Coururier, A.; Le Guellec, K.; Omilli, F.; Camonis, J.; Mac
Proc. Natl. Acad. Sci. U.S.A. 88, 1039-1043, 1991
A:Title: Cloning by differential screening of a Xenopus cDNA coding for a protein highly
A:Reference number: A37871; MUID:91126051; PMID:1704128

A:Accession: A37871
A:Molecule type: mRNA
A:Residues: 1-297 <PAR>
A:Cross-references: GB:X14227
R:Le Guellec, R. EMBL Data Library, January 1989
submitted to the EMBL Data Library, January 1989
A:Reference number: S15866
A:Accession: S15866
A:Molecule type: mRNA
A:Residues: 1-92, 'R', '94-297 <LEA>
A:Cross-references: EMBL:X14227; NID:g64665; PIDN:CAA32443.1; PID:g64666
R:Olive, M.; Theze, N.; Philippe, M.; Le Pennec, J.P.; Lerivray, H.
Gene 151, 81-88, 1994
A:Title: Cloning of the Xenopus laevis cdk2 promoter and functional analysis in oocyte
A:Reference number: I51662; MUID:95129896; PMID:7828909
A:Accession: I51662
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-39 <OLI>
A:Cross-references: EMBL:U07979; NID:g473584; PIDN:AAA82123.1; PID:g473585
C:Genetics:
A:Gene: cdk2
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.5%; Score 555; DB 2; Length 297;
Best Local Similarity 37.8%; Pred. No. 5.9e-19;
Matches 115; Conservative 59; Mismatches 90; Indels 40; Gaps 6;

QY 1 MEKYELKATGEGSGVGVFKCNKTSQGVAVKFESEDDPVVKKIALREIMLKQLKH 60
DB 1 MEKYEKELKATGEGSGVGVFKCNKTSQGVAVKFESEDDPVVKKIALREIMLKQLKH 60
QY 61 PNLVNLIEFRKRKMHVLFVEYCDHTLLNELE-RNPNGVADGVKSVLWQTLQALNFCIH 119
DB 61 PNLVNLIEFRKRKMHVLFVEYCDHTLLNELE-RNPNGVADGVKSVLWQTLQALNFCIH 119
QY 120 NCIHRIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDVYATRWYRAPELLVGDQYQ 178
DB 120 NCIHRIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDVYATRWYRAPELLVGDQYQ 178
QY 121 RVLHRDLKPNQLLNSDCAIKLADFGARAFGVPVRYTSHEVVTLYWYRAPELLVGDQYQ 180
DB 121 RVLHRDLKPNQLLNSDCAIKLADFGARAFGVPVRYTSHEVVTLYWYRAPELLVGDQYQ 180
QY 179 GSVDVMAIGCVFAELLTGQPLWPKSDVDQVLYLIRTL----- 217
DB 179 GSVDVMAIGCVFAELLTGQPLWPKSDVDQVLYLIRTL----- 217
QY 181 STAVDWSLGCIFAETFRALFGDSEIDQVFRIFRILGTPDEVSWPGVTMPDYKSTF 240
DB 181 STAVDWSLGCIFAETFRALFGDSEIDQVFRIFRILGTPDEVSWPGVTMPDYKSTF 240
QY 218 ---VETGFRHV----DQAGLELLTSDDPFAVASOSAGITGKLIPRHQSIFKSNQFFHGIS 270
DB 218 ---VETGFRHV----DQAGLELLTSDDPFAVASOSAGITGKLIPRHQSIFKSNQFFHGIS 270
QY 241 PKWIRCFPSKVVVPLDEDGRDLAQM---LQYDSNKRISAKVALTHP-----FFRDVS 290
DB 241 PKWIRCFPSKVVVPLDEDGRDLAQM---LQYDSNKRISAKVALTHP-----FFRDVS 290
QY 271 IPEP 274
DB 271 IPEP 274
QY 291 RPTP 294
DB 291 RPTP 294

RESULT 7

A41227
Protein kinase (EC 2.7.1.37) cdk2 - human
N:Alternate names: Egl homolog; protein kinase p34
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
A:Accession: A41227; S17873; S16520
R:Ninomiya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A:Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of
A:Reference number: A41227; MUID:92020980; PMID:1717994
A:Accession: A41227
A:Molecule type: mRNA
A:Residues: 1-298 <NIN>
A:Cross-references: GB:M69520; NID:gl80177; PIDN:AAA35667.1; PID:gl80178
R:Tsay, L.H.; Harlow, E.; Meyerson, M.

C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.4%; Score 552.5; DB 2; Length 294;
Best Local Similarity 35.4%; Pred. No. 7.5e-19;
Matches 119; Conservative 66; Mismatches 98; Indels 53; Gaps 9;

QY 1 MEKYEKLAKTGEGSGVGVVCKRNKTSQGVAVKKFVESEDDPVVKIALREYRMLKQLKH 60
DB 1 MEQYEVKEKIGSGYGVVYKALDKTANETALKKIRLEQDEGCVSTAIRSELILKEVNH 60
QY 61 PNLVNLIVFPRKAKMHLFVEYCDHTLNL-----ELERNPNQVADGVTKSVLWOTLQAL 114
DB 61 GNIVRLHDVHASEKRIYLFVEYLDLKKFMDSCPFAKNPT-----LIKSYLYQLIRGV 115
QY 115 NCFTHNCIHROIKPENILITKQ-GIIKICDFGFAOIL-LPGDAYTVVATRYRAPELL 172
DB 116 AYCHSRVLRDLAPQNLIDRRNNAUKADFGARAFGIPVTRTFHEVVTULWYRAPELL 175
QY 173 VGDQYSGSDVIAIGCVFAELLTGQPLNFGKSDVDQLYLIINTL---VETQFRHWDAQ 229
DB 176 LCAQYSTPVDVMSVGCIFAEVMVQKFLPFGDSEIDELFKIFVLTGFEQGW----- 228
QY 230 LELLTSSDPPAVASQSGITCKLIPRHOSIFKSNGFPHGISIPEPEMDMETLSEKSDVHP 289
DB 229 -----PGVSLDPKFA--FPRWQA-----QDLATI---VPLLEP 258
QY 290 VALNPMKCLXNPDRLTCSQLLESSYFDSFQEAQ 325
DB 259 AGLDLSLKNLRYEPSKRITARQALEHEYFKDLEWVQ 294

RESULT 3
S23095
protein kinase (EC 2.7.1.37) cdc2 - Arabidopsis thaliana
N;Alternate names: cdc2 protein homolog; CDC2a protein; cell division control protein 2
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #ext_change 20-Jun-2000
C;Accession: S23095; A48984; JQ1337; JQ0967; S18202
R;Imajuku, Y.; Hirayama, T.; Endoh, H.; Oka, A.
FEBS Lett. 304, 73-77, 1992
A;Title: Exon-intron organization of the Arabidopsis thaliana protein kinase genes CDC2
A;Reference number: S23095; MUID:92316202; PMID:1618302
A;Accession: S23095
A;Molecule type: DNA
A;Residues: 1-294 <IN>
A;Cross-references: EMBL:D10850; NID:g217848; PiDN:BA01623.1; PID:g217849
R;Inze, D.; Ferreira, P.; Hemerly, A.; Van Montagu, M.
Biochem. Soc. Trans. 20, 80-84, 1992
A;Title: Control of cell division in plants.
A;Reference number: A48984; MUID:92339744; PMID:1634002
A;Accession: A48984
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-294 <INZ>
A;Experimental source: flower
A;Note: Sequence extracted from NCBI backbone (NCBI:P:109461)
R;Hirayama, T.; Imajuku, Y.; Anai, T.; Matsui, M.; Oka, A.
Gene 105, 159-165, 1991
A;Title: Identification of two cell-cycle-controlling cdc2 gene homologs in Arabidopsis
A;Reference number: JQ1337; MUID:92039027; PMID:1937013
A;Accession: JQ1337
A;Molecule type: mRNA
A;Residues: 1-294 <HIP>
A;Cross-references: EMBL:X57839; NID:q16218; PiDN:CAA40971.1; PID:g16219
R;Ferreira, P.C.G.; Hemerly, A.S.; Villarroel, R.; Van Montagu, M.; Inze, D.
Plant Cell 3, 531-540, 1991
A;Title: The Arabidopsis functional homolog of the p34cdc2 protein kinase.
A;Reference number: JQ0967; MUID:93005715; PMID:1840925
A;Accession: JQ0967

```
A:Molecule type: mRNA
A:Residues: 1-294 <FER>
A:Cross-references: GB:S45387; NID:G257373; PIDN:AA321643.1; PID:G257374
A:Experimental source: flower
C:Comment: The protein is a key component of the eukaryotic cell cycle.
C:Genetics:
A:Gene: cdc2
A:Introns: 3/3; 65/2; 105/3; 163/3; 218/2; 247/3; 265/3
C:Function:
A:Description: phosphotransferase; protein kinase; required for G1 to S-phase transition
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33.51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 35.5%; Pred. No. 8.4e-19;
Matches 115; Conservative 71; Mismatches 95; Indels 43; Gaps 8;

QY 1 MEKYEKAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVVYKIALREIRMLKQLKH 60
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MDQYEKKEIGEGTGVVYKARDKVTNETIALKIRLEQDEGVSTAIRERISLLKEMQH 60

QY 61 PNLVNLIEVFRKRKHILVFYCDHTLLNLELRPNGVAD-GVLSVLMOTLQALNECHI 119
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 SNIVKLDVHSEKRLYLVEYLDLDLKKHMDSTPFSKDLHMKTLYQLRGIAVCHS 120

QY 120 HNCIHRDIKPNILITKO-GIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDQ 177
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 HRVLRDLKPNQLLIDRTNSLKLADFLGARAFGIPVRTFTHEVVTWYRAPEILLGSH 180

QY 178 YGSSVDIWAICGVFAELLTGCPWPKGSDVDOLYLIR---TLVETGFRHVDQAGLELL 234
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 YSTPVD:MSVGCIFAEMLISQPLPFGDSEIDPLKPIRIMGTPYEDTWR----- 229

QY 235 SSDPPAVASQAGITGLIPRHQIFKSNFGFHGISPEPEDEMETLEEKFSVDHPVALNF 294
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
230 -----GVTS--LPDYKSAFPK-----WKPTDLETF---VPLDPCVDL 263

QY 295 MKGCLXNPPDRJTCSSQLLESSYF 318
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 LSKMLMDPTKRIARAALAEHYF 287

RESULT 10
S22440
Protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice
C:Species: Oryza sativa (rice);
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S22440
R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuki, I.; Utsugi,
Mol. Gen. Genet. 233, 10-16, 1992
A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory
A:Reference number: S22440; MUID:92293101; PMID:1376401
A:Accession: S22440
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <HAS>
A:Cross-references: ENBL:X60374; NID:G20343; PIDN:CAA42922.1; PID:Q20343
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33.51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 34.8%; Pred. No. 8.4e-19;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;

QY 1 MEKYEKAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVVYKIALREIRMLKQLKH 60
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MEQYEKKEIGEGTGVVYKARDKVTNETIALKIRLEQDEGVSTAIRERISLLKEMQH 60

RESULT 11
A40444
Protein kinase (EC 2.7.1.37) cdc2 homolog A - maize
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A40444
R:Colasanti, J.; Tyers, M.; Sundaresan, V.
Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991
A:Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2)
A:Reference number: A40444; MUID:91195354; PMID:2014258
A:Accession: A40444
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-294 <COL>
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33.51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.0%; Score 546.5; DB 1; Length 294;
Best Local Similarity 34.1%; Pred. No. 1.4e-18;
Matches 117; Conservative 66; Mismatches 93; Indels 67; Gaps 9;

QY 1 MEKYEKAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVVYKIALREIRMLKQLKH 60
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MEQYEKKEIGEGTGVVYKALDTANETIALKIRLEQDEGVSTAIRERISLLKEMNH 60

QY 61 PNLVNLIEVFRKRKHILVFYCDHTLLN-----ELERNPNGVADGVLSVLMOTLQAL 114
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 GNIVLRHDVHSEKRLYLVEYLDLDLKKFMDSCPEFAKNPT-----LIKSYLYQLRGV 115

QY 115 NFCHNCIHRDIKPNILITKO-GIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELL 172
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 AYCHSHRVLRDLKPNQLLIDRTNALKLADFLGARAFGIPVRTFTHEVVTWYRAPEIL 175

QY 173 VGDYTGSSVDIWAICGVFAELLTGCPWPKGSDVDOLYLIRLTVETGFRHVDQAGLEL 232
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 LGSROYSTPVDVMSVGCFAEWMNKPLPFGDSEIDELFKIPVLGTP----- 223

QY 233 LTSDDPPAVASQAGITGKLIPRHQIFKSNFGFHGISPEPEDEMETLEEKFSVDHPVAL 292
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 -NEQSWPGVSS-----LPDYKSAFPK-----WQAQDLATI---VPTLDPAGL 261

QY 293 NFMKGLXNPPDRJTCSSQLLESSYFDSFQEAQ 325
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
262 DLSKMKRYEPKRIARQAALAEHYFKDLEMVQ 294

RESULT 12
A40444
Protein kinase (EC 2.7.1.37) cdc2 homolog A - maize
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A40444
R:Colasanti, J.; Tyers, M.; Sundaresan, V.
Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991
A:Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2)
A:Reference number: A40444; MUID:91195354; PMID:2014258
A:Accession: A40444
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-294 <COL>
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33.51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.0%; Score 546.5; DB 1; Length 294;
Best Local Similarity 34.1%; Pred. No. 1.4e-18;
Matches 117; Conservative 66; Mismatches 93; Indels 67; Gaps 9;

QY 1 MEKYEKAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVVYKIALREIRMLKQLKH 60
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MEQYEKKEIGEGTGVVYKALDTANETIALKIRLEQDEGVSTAIRERISLLKEMNH 60

QY 61 PNLVNLIEVFRKRKHILVFYCDHTLLN-----ELERNPNGVADGVLSVLMOTLQAL 114
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 GNIVLRHDVHSEKRLYLVEYLDLDLKKFMDSCPEFAKNPT-----LIKSYLYQLRGV 115

QY 115 NFCHNCIHRDIKPNILITKO-GIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELL 172
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 AYCHSHRVLRDLKPNQLLIDRTNALKLADFLGARAFGIPVRTFTHEVVTWYRAPEIL 175

QY 173 VGDYTGSSVDIWAICGVFAELLTGCPWPKGSDVDOLYLIRLTVETGFRHVDQAGLEL 232
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 LGSROYSTPVDVMSVGCFAEWMNKPLPFGDSEIDELFKIPVLGTP----- 221

QY 233 LTSDDPPAVASQAGITGKLIPRHQIFKSNFGFHGIS-IP-----EPEDMETLEE 282
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 -----TPNEQS-----WPGVSCLPDFKTAFFRMAQDLATV-- 252

QY 283 KFSVDHPVALNFMKGLXNPPDRJTCSSQLLESSYFDSFQEAQ 325
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 -VPLDPCVDLSSMLRYEPKRIARQAALAEHYFKDLEVVQ 294

RESULT 12
```

T49271
CELL DIVISION CONTROL PROTEIN 2 HOMOLOG A - Arabidopsis thaliana
N:Alternate names: protein T49271.20
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000
C:Accession: T49271
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Rudd, S.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25021
A:Accession: T49271
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <RIF>
A:Cross-references: EMBL:AL132963; GSPDB:GN0006; ATSP:T21J18.20
A:Experimental source: cultivar Columbia; BAC clone T21J18
C:Genetics:
A:Gene: ATSP:T21J18.20
A:Map position: 3
A:Introns: 3/3; 67/3; 105/3; 163/3; 219/2; 247/3; 265/3
C:Superfamily: kinase-related transforming protein; protein kinase homology
Query Match 30.0%; Score 546.5; DB 2; Length 294;
Best Local Similarity 35.2%; Pred. No. 1.4e-18;
Matches 1:4; Conservative 7; Mismatches 96; Indels 43; Gaps 8;
QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLK 60
DB 1 MDQYKVEKIGEGTYGVVYKARDKVTNETIALKKIRLEQDEGVSTAIRISLLKEXCH 60
QY 61 PNVLNLIIEVFRKRKXHLVFEYCDHTLLNELERNENG--VADGVKSVLWQTLQALNFC 119
DB 61 SNIVYDVTSEKRLYLVFEYLDJDLKEMDSTPSSKOLHKIKTYLQILRGIAVCS 120
QY 120 HNCIHRDKPENILITKQGIKICDFGFAQL-IPGDAYTYVATRYRAPELVGDTQ 177
DB 121 HRVLHRLKPKQNLIDRTNLSKLADFLARAFGIPVTRFHEVVLYWYRAPELVGSHH 180
QY 178 YGSSVDIWAIGCVFAELLTGQPLWPKGSDVQQLIIR---TLVETGRVVDQAGLELIT 214
DB 181 YSTPVDIWSVGCFAEMISQKLEFPGSDSEIDQLFKIFRMGTPTVEDTWR----- 229
QY 235 SSDPPAVASQAGITGKLIPIHQSGIFKSNGFHGISIPEDMETLEEKSDVHPVALNF 234
DB 230 -----GVTS--LPDYKSAFPK-----WKPTDLETF---VFNLDPPGVDL 263
QY 295 MKGCLAKRPDRLTCSQLSESYF 318
DB 264 LSKMLMDPTKRNARAAJSEHYF 287
RESULT 13
A29539
protein kinase (EC 2.7.1.37) cdc2 - human
N:Alternate names: cell division control protein 2 (CDC2)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jul-2000
C:Accession: A29539
R:Lee, M.G.; Nurse, P.
Nature 327, 31-35, 1987
A:Title: Complementation used to clone a human homologue of the fission yeast cell cycle
A:Reference number: A29539; MUID:8720:915; PMID:3553962
A:Accession: A29539
A:Molecule type: mRNA
A:Residues: 1-297 <LEE>
A:Cross-references: GB:X05360; NID:929838; PIDN:CAA28963.1; PID:929839
C:Genetics:
A:Gene: GDB:CDC2
A:Cross-references: GDB:119052; OMIM:116940
A:Map position: 10q21.1-10q21.1
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif

F:14-161/Binding site: phosphate (Thr) (covalent) #status predicted
F:15/Binding site: phosphate (Cys, Glu, Asp, Lys #status predicted
F:33:51,128,132/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 29.9%; Score 545; DB 2; Length 297;
Best Local Similarity 40.2%; Pred. No. 1.7e-18;
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;
QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLK 59
DB 1 MEDYTKIEKIGEGTYGVVYGRHKTKTQGVVAMKKIRLESEEG-VFSTAIRISLLKELR 59
QY 60 HPNLVNLIEVFRKRKXHLVFEYCDHTLLNELERNENG--VADGVKSVLWQTLQALNFC 117
DB 60 HPNIVSLQDVLMOQSRLYLIFEFLSMDLKKYLDLSPGQYMDSSLVSKSYLQQLQGVFC 119
QY 118 HNCIHRDKPENILITKQGIKICDFGFAQL-IPGDAYTYVATRYRAPELVGDT 176
DB 120 HSRVLHRLKPKQNLIDRTNLSKLADFLARAFGIPRVYTHVVTWYRSPEVLGSA 179
QY 177 YGSSVDIWAIGCVFAELLTGQPLWPKGSDVQQLIIRTL----- 217
DB 180 RYSTPVDIWSIGTIFAEALAKKPLFGDSEIDQLFRIFRALGTFNNEVWPEVESLQDYKN 239
QY 218 -----VETGRVVDQAGLELITSS---DPPAVASQAGITGKLIPIRH 256
DB 240 TFPKWKPGSLASHVKNLDEGLLSKMLYDP-----AKRISGRMAJNH 284
RESULT 14
A44878
protein kinase (EC 2.7.1.37) cdk2 (validated) - goldfish
C:Species: Carassius auratus (goldfish)
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: A44878
R:Hi-rai, T.; Yamaehita, M.; Yoshikuni, M.; Tokumoto, T.; Kajiuira, H.; Sakai, N.; Nagahai
Dev. Biol. 152, 113-120, 1992
A:Title: Isolation and characterization of goldfish cdk2, a cognate variant of the cell
A:Reference number: A44878; MUID:92331802; PMID:1339336
A:Accession: A44878
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-298 <HIR>
A:Cross-references: GB:S40289; NID:9251619; PIDN:AAB22550.1; PID:9251620
A:Experimental source: oocyte
A:Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBI:P:108783)
C:Genetics:
A:Gene: cdk2
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serin
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33:51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 29.9%; Score 545; DB 1; Length 298;
Best Local Similarity 46.6%; Pred. No. 1.7e-18;
Matches 102; Conservative 48; Mismatches 67; Indels 2; Gaps 2;
QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLK 60
DB 1 MESQFKVEKIGEGTYGVVYKAKKKTGETVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNVLNLIIEVFRKRKXHLVFEYCDHTLLNELERNP--NGVADGVKSVLWQTLQALNFC 119
DB 61 PNVLHLDVHTENKYLVEFLHQDLKRFMDSTVTGSLPLVSKSYLQQLQGLAFCHS 120
QY 120 HNCIHRDKPENILITKQGIKICDFGFAQL-IPGDAYTYVATRYRAPELVGDTQY 178
DB 121 HRVLHRLKPKQNLIDRTNLSKLADFLARAFGIPVTRFHEVVLYWYRAPELVGCKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPKGSDVQQLIIRTL 217
DB 181 STAVDIWSLGCIFAEMITRKALFPDSEIDQLFRIFRTL 219

```

RESULT 15
I45977
cyclin-dependent kinase 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 18-Jun-1999
C:Accession: I45977
R:Yang, L.; Fatin, C.E.
Gene 141, 283-286, 1994
A:Title: Identification of cDNAs encoding bovine cyclin B and Cdk1/Cdc2.
A:Reference number: I45977; XUID:9421918; PMID:8183283
A:Accession: I45977
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-297 <YAN>
A:Cross-references: GB:L26547; NID:G433155; PIDN:AAA18894.1; PID:G498-73
C:Genetics:
A:Gene: cdk1/cdc2
C:Superfamily: Kinase-related transforming protein; protein kinase homology
F:2-256/Domain: protein kinase homology <KIN>

Query Match          29.8%; Score 542; DB 2; Length 297;
Best Local Similarity 40.2%; Pred. No. 2.3e-18;
Matches 117; Conservative 57; Mismatches 75; Indels 42; Gaps 7;

QY 1 MEKYELAKTGEQYGVVFKCRNKTSGOVAVKKP-VESEDDPVVKKIALRE:RMLKQLK 59
DB 1 MEDYTKIEIGEGTYGVVYKGRHKTTGOVAMKKIRLESEEG-VPSTAIRISLLKELR 59

QY 60 HPNLVNLIEVFRKKRMHLVFEYCDHTLLNELRNPG--VAGGVKSVLWQTLQALNFC 117
DB 60 HPNTVSLQDVLMQDSRLYLIFELSLMDLKKYLDSPGQFMDSSEVKSYS:QILQGIIVFC 119

QY 118 HHHNCIHRDIKPNILITKQGIHKICDPGFAQIL-IFGDAYTDYVATRWYRAPELLVGET 176
DB 120 HRRVLRDLKPONLLIDDKGTIKLADFLARAFGIQ:RVYTHEWTLKYRSPEVLGSA 179

QY 177 QYGSVDIWAIGCVFAEELLTQCLPMPGKSDVDQLYLIIRTL----- 217
DB 180 GYSTPVDIWS:GTIFAEALTKKPLFHGDSIDQLFRIFRALGTFPNEWPEVESLDYKS 239

QY 218 -----VETGRHVDQAGLELLTSS---DPPAVASQASGITGKLIPIRH 256
DB 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYCP-----AKPISGRKVALNH 284

```

Search completed: November 13, 2003, 14:21:16
Job time : 15.6767 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein: search, using sw model

Run on: November 13, 2003, 14:12:56 ; Search time 10.4834 Seconds
(without alignments)
1556.584 Million cell updates/sec

Title: US-09-671-050-10

Perfect score: 1820

Sequence: 1 MEKYEKIAKTGEGSYGVVFX.....RXARNEGRNRROVLEPLKS 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1197.5	63.6	358	1	Q00532 homo sapien
2	694.5	38.2	1030	1	Q76039 homo sapien
3	565	31.0	305	1	Q00526 homo sapien
4	563	30.9	302	1	P51958 carassius a
5	561	30.8	297	1	P34337 xenopus lae
6	558.5	30.7	292	1	P34117 dictyostell
7	554	30.4	298	1	P24918 homo sapien
8	551.5	30.3	294	1	P24100 arabidopsis
9	551.5	30.3	294	1	Q85076 cricetus
10	550	30.2	298	1	Q8dga2 oryzias jav
11	549.5	30.2	303	1	Q8dga2 oryzias jav
12	548	30.1	296	1	Q8dga2 oryzias jav
13	546	30.0	298	1	Q8dga2 oryzias jav
14	545	29.9	297	1	Q8dga2 oryzias jav
15	545	29.9	298	1	Q8dga2 oryzias jav
16	543.5	29.9	303	1	Q8dga2 oryzias lat
17	542.5	29.8	294	1	Q8dga2 oryzias lat
18	542	29.8	297	1	Q8dga2 oryzias lat
19	541.5	29.8	294	1	Q8dga2 oryzias lat
20	541.5	29.8	303	1	Q8dga2 oryzias lat
21	541.5	29.8	303	1	Q8dga2 oryzias lat
22	541	29.7	297	1	Q8dga2 oryzias lat
23	540.5	29.7	297	1	Q8dga2 oryzias lat
24	540.5	29.7	302	1	Q8dga2 oryzias lat
25	540	29.7	302	1	Q8dga2 oryzias lat
26	539	29.6	303	1	Q8dga2 oryzias lat
27	537	29.5	288	1	Q8dga2 oryzias lat
28	536.5	29.5	294	1	Q8dga2 oryzias lat
29	536.5	29.5	311	1	Q8dga2 oryzias lat
30	535.5	29.4	294	1	Q8dga2 oryzias lat
31	532	29.2	297	1	Q8dga2 oryzias lat
32	532	29.2	301	1	Q8dga2 oryzias lat
33	531	29.2	292	1	Q8dga2 oryzias lat

RESULT 1

KKIA_HUMAN 34 531 29.2 292 1 CDK5 RAT Q03114 rattus norv
35 530.5 29.1 294 1 CC22 MEDSA Q05006 medicago sa
36 529 29.1 292 1 CDK5 MOUSE P49615 mus musculu
37 527 29.0 301 1 CC21 TRYCO P54664 trypanosoma
38 526 28.9 292 1 CDK5_HUMAN Q00535 homo sapien
39 526 28.9 296 1 CDK2_DCDI P34112 dictyostell
40 525 28.8 294 1 CDK5 DROME P48609 drosophila
41 521.5 28.7 294 1 CC2A ANTMA Q38772 antirrhinum
42 521 28.6 292 1 CDK5_XENLA P51166 xenopus lae
43 519.5 28.5 346 1 CDK7_HUMAN P50613 homo sapien
44 518 28.5 292 1 CC22_OEYSA P29619 oryza sativ
45 518 28.5 305 1 PH85_YEAST P17157 saccharomyc

ALIGNMENTS

ID ID KKIA_HUMAN STANDARD; PRT; 358 AA.
AC Q00532;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase KKIALRE (EC 2.7.1.-) (Cyclin-dependent
DE kinase-like 1).
GN CDK5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=92347325; PubMed=1639063;
RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,
RA Harlow E., Tsai L.-H.;
RT "A family of human cdc2-related protein kinases."
RL EXBO J. 11:2909-2917(1992).
CL -!
CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X66358; CAA47002.1; -;
CC EVBL: X66359; CAA47002.1; JOINED.
CC FJR: S23383; S23383.
CC HSP: P24941; LHCL.
CC Genew; HGNC:1781; CDKL1.
CC GK: Q00532; -;
CC XIM: 603441; -;
CC GO: GO:0004693; F:cyclin-dependent protein kinase activity; TAS.
CC GO: GO:0004688; P:protein amino acid phosphorylation; TAS.
CC GO: GO:0000074; P:regulation of cell cycle; TAS.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR002290; Ser_thr_kinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TKc_1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferase: Serine/threonine-protein kinase: ATP-binding.
CC DOMAIN 5 288
CC NP_BIND 11 19
CC BINDING 34 34
CC ACT_SITE 127 127
CC SEQUENCE 358 AA: 41834 MW: 88344321F24B77C6 CRC64;

```
Query Match 63.6%; Score 1157.5; DB 1; Length 358;
Best Local Similarity 64.7%; Pred. No. 9.8e-75;
Matches 211; Conservative 41; Mismatches 41; Indels 33; Gaps 2;

QY 1 MEKYEKLAKTGEQGVVFKCRNKTSGQVAVKVFSEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGEQGVVFKCRNKTSGQVAVKVFSEDDPVVKKIALREIRMLKQLKH 61
QY 61 PNLVNLIEVFRKRKHLVFEYCDHTLLNELRPNVGVADGVKSVLWOTLQALNFCIH 120
DB 62 PNLVNLIEVFRKRKHLVFEYCDHTLLNELRPNVGVADGVKSVLWOTLQALNFCIH 121
QY 121 NCIHRIKPNILITKQGIKICDFGAQILI-PGDAYTDYVATRYAPPELLVGDQYQ 179
DB 122 NCIHRIKPNILITKQGIKICDFGAQILI-PGDAYTDYVATRYAPPELLVGDQYQ 181
QY 180 SSVDLVWAGCVFAELLTGQPLWPKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDP 239
DB 182 PPDVDMVWAGCVFAELLTGQPLWPKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDP 239
QY 240 AVASQSGAGITGKLIHQKIFKSGFFRGIS-PEPEMETLEEKFSVHVPVAVLWYKGC 299
DB 220 -----GDLIPRHQQVSTNCFYSGVKIPDPEDMEPIELKFPNISYFALGLLKGCL 269
QY 300 KNPDPDLRLTCSQLLESSYFDSFQBAQ 325
DB 270 HMDPTERTLCQLLHHPYFENIREIE 295

RESULT 2
STK9_HUMAN
ID STK9_HUMAN STANDARD; PRT; 1030 AA.
AC O76039; Q4199;
DT 15-JUL-1999 (Rel. 38, Created)
DE 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein kinase 9 (EC 2.7.1.37).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=98389626; PubMed=9721213;
RA Montini E., Andolfi G., Caruso A., Buchner G., Walpole S.M.,
RA Mariani M., Consalez G.G., Trimp D., Ballabio A., Franco B.;
RT "Identification and characterization of a novel serine-threonine
RT kinase gene from the Xp22 region.";
RL Genomics 51:427-433(1998).
RN [2]
RP SEQUENCE OF 339-789 FROM N.A.
RA Krause S.W., Reili M., Kreutz M., Schwarzfischer L., Paulauskis J.L.,
RA Andreesen J.D.;
RT "Differential screening leads to novel genetic markers of monocyte
RT to macrophage maturation.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP -> ADP + a protein = ADP - a phosphoprotein.
CC -!- ENZYME REGULATION: COULD BE ACTIVATED BY PHOSPHORYLATION ON
CC TYROSINE AND THREONINE. PROBABLY BOTH PHOSPHORYLATIONS ARE
CC REQUIRED FOR ACTIVITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, KIDNEY, PROSTATE,
CC OVARY, PLACENTA, PANCREAS AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 415.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; Y15057; CAA75342.1; -;
EMBL; X89059; CAA61445.1; ALT_FRAME.
HSSP; P24941; 1B38.
Genes; HGNC:11411; STK9.
GK; O76039; -;
MIM; 300203; -;
GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
GO; GO:0005648; P:protein amino acid phosphorylation; TAS.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; pkinae.1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TK; 1.
PROSITE; PS0107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
Transferrase; Serine/threonine-protein kinase; ATP-binding;
Phosphorylation.
DOMAIN 13 297 PROTEIN KINASE.
NP_BIND 19 27 ATP (BY SIMILARITY).
BINDING 42 42 ATP (BY SIMILARITY).
ACT_SITE 135 135 BY SIMILARITY.
MOD_RES 169 169 PHOSPHORYLATION (ACTIVATES THE KINASE)
(POTENTIAL).
MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
(POTENTIAL).
DOMAIN 784 789 POLY-LYS.
CONFLICT 339 340 HR -> GT (IN REF. 2).
CONFLICT 541 541 L -> W (IN REF. 2).
CONFLICT 731 764 MISSING (IN REF. 2).
SEQUENCE 1030 AA; 115537 MW; 6A1C9C438610EF08 CRC64;

Query Match 38.2%; Score 694.5; DB 1; Length 1030;
Best Local Similarity 40.3%; Pred. No. 1.5e-41;
Matches 139; Conservative 69; Mismatches 98; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGEQGVVFKCRNKTSGQVAVKVFSEDDPVVKKIALREIRMLKQLKH 60
DB 10 KNPPELLGVGEGAYGVVLCRKHETHEIIVAIFKDSSENEEVKETTLELQKRLTKQ 69
QY 61 PNLVNLIEVFRKRKHLVFEYCDHTLLNELRPNVGVADGVKSVLWOTLQALNFCIH 120
DB 70 ENIVELKEAPRRCKLYLVFEYVEKMLELLEEMPNVGPPEKVKSYIYQLIKAIHWCKN 129
QY 121 NCIHRIKPNILITKQGIKICDFGAQILI-PGDAYTDYVATRYAPPELLVGDQYQ 178
DB 130 DIVHRDIKPNILLISHNDVLKLCDFGFARNLSENNANYTEYVATRYAPPELLVGDQYQ 188
QY 179 GSSVDVWAGCVFAELLTGQPLWPKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDP 238
DB 189 GKSVDVWAGCVFAELLTGQPLWPKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDP 227
QY 239 PAVASQSGAGITGKLIHQKIFKSGFFRGIS-PEPEMETLEEKFSVHVPVAVLWYKGC 297
DB 228 -----GDLIPRHQQVSTNCFYSGVKIPDPEDMEPIELKFPNISYFALGLLKGCL 276
QY 298 CLKNPDPDLRLTCSQLLESSYFDSFQBAQIKRKARNEGNNRRQQV 342
DB 277 LLKLDPADRYLTYECLN---HPTQTORILLDRSPRSKAKPKYHV 318

RESULT 3
CDK3_HUMAN
ID CDK3_HUMAN STANDARD; PRT; 305 AA.
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
```


Cell division protein kinase 3 (EC 2.7.1.-).
 GN CDK3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN 1;
 RF TISSUE=Fetal brain;
 RC MEDLINE=92347325; PubMed=1639063;
 RA Meyerson M., Enders G.H., Wu C.-L., Su L.-X., Gorka C., Nelson C.,
 RA Harlow E., Tsai L.-H.;
 RT "A family of human cdc2-related protein kinases.";
 RL EMBO J. 11:2909-2917(1992).
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
 CC INTERACTS WITH A YET UNKNOWN TYPE OF CYCLIN. CAN PHOSPHORYLATE
 CC HISTONE H1.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC2/CDKX SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X66357; CAA47001.1; -.
 CC PIR: S23382; S23382.
 CC PDB: 1LFN; 24-APR-02.
 CC Genew: HGNC:1772; CDK3.
 CC GK: Q00526; -.
 CC MIM: 123828; -.
 CC GO: GO:0004693; P:cyclin-dependent protein kinase activity; TAS.
 CC GO: GO:0007048; P:oncogenesis; TAS.
 CC GO: GO:0000074; P:regulation of cell cycle; TAS.
 CC InterPro: IPR000719; Prot kinase.
 CC InterPro: IPR002290; Ser/thr_pkinase.
 CC Pfam: PF00069; pkinase; 1.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC SMART: SM00220; S_TKc; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 CC Transferase: Serine/threonine-protein kinase; ATP-binding; Cell cycle;
 CC Cell division; Mitosis; Phosphorylation; 3D-structure.
 CC DOMAIN 4 286 PROTEIN_KINASE
 CC FT NP_BIND 10 18 ATP (BY SIMILARITY).
 CC FT BINDING 33 33 ATP (BY SIMILARITY).
 CC FT ACT_SITE 127 127 BY SIMILARITY.
 CC FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 160 160 PHOSPHORYLATION (BY SIMILARITY).
 CC SEQUENCE 305 AA; 35045 MW; 1128BE0C96EA262A CRC64;
 Query Match 31.0%; Score 565, DB 1; Length 305;
 Best Local Similarity 36.8%; Pred. No. 5e-33;
 Matches 124; Conservative 66; Mismatches 99; Indels 48; Gaps 3;
 QY 1 MEKYEKLATGEGVGWFKCRKTSQGVAVKFE---VESEDDPVWKKIALREIPMLKC 57
 DB 1 MDVQKVEKIGETGYGVYKAKNRRTGQVAVLKKTLDLEKGV---STAIRSILKE 57
 QY 58 LKHPNLVNLGIEVFRKRMKHLFEYCDHTLLNLELRNPNG-VADGVKSVLWQTLQALNF 116
 DB 58 LKHPNLVRLDDVHNRKLVLYFEF--SQDLKKYMDSTPGSELPLKIKSYLFOLLQGVSF 117
 QY 117 CHTHNCIHRDIPKENVLTZKGGIHKICDFGFAQL-IPCDATDVATKRYRPELLVGD 175
 DB 118 CHSHRVHDLKPKQNLINELGAIKLADGLARAFGVPLRTYTHEWTLVYRAPEILGS 177
 QY 176 TQYGSSVD-WAIGCVFAELLTGPIWPGKSDVDQLYLIRTVETGFRHVDQAGLEILTS 235

"The isolation from a unicellular organism, Dictyostelium discoideum, of a highly-related cdc2 gene with characteristics of the PCNAIRE subfamily."

RL Biochim. Biophys. Acta 1179:117-124(1993).

CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES

CC THE ENZYME (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDC2/CDKX SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: J00652; AAA16056.1; -

DR FIR: S40021; S40021.

DR HSSP: P24941; 1CRP.

DR DictyDB: D205039; crpA.

DR InterPro: IPR000719; Prot Kinase.

DR InterPro: IPR002290; Ser Thr kinase.

DR Pfam: PF00069; pkinase; 1

DR ProDom: PD000001; Prot_kinase; 1.

DR SMART: SM00220; S_TKc; 1.

DR PROSITE: P500107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: P500108; PROTEIN_KINASE_ST; 1.

DR PROSITE: P50011; PROTEIN_KINASE_DOM; 1.

KW Transferase: Serine/threonine-protein kinase; ATP-binding;

KW Phosphorylation.

FT DOMAIN 4 285 PROTEIN KINASE.

FT NP_BIND 10 18 ATP (BY SIMILARITY).

FT BINDING 33 33 BY SIMILARITY).

FT ACT_SITE 126 126 BY SIMILARITY.

FT MOD_RES 14 15 PHOSPHORYLATION (BY SIMILARITY).

FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).

SQ SEQUENCE 292 AA; 33226 MW; 361AB54C4E3BD41E CRC64;

Query Match: 30.7%; Score 558.5; DB 1; Length 292;

Best Local Similarity 35.0%; Pred. No. 1.4e-32;

Matches 155; Conservative 66; Mismatches 10; Indels 47; Gaps 6;

QY 1 MEKYELAKTGGSGVGFCKNKTSGGVAVKFKVESEDPVVKKIALREIRLXKQ-KY 60

Db 1 MEKYKIEKLGSTGVYVNAKAKRTGETGEIVALKRILDSSEGVPTAIRISLLKELKH 60

QY 61 PHLVNLIEVFRKRRKHLVFEYCDHTLLNLEERNPNVGVKISVLQMOTLQALNPFCHH 120

Db 61 PHLVNLVHIVHTEKLLTVFEYLDQDLKKYLDCEGGEISKPTIKSFYQLKGVAFCHDH 120

QY 121 NCIRHGIKPNILITKGIKICDFGFAQIL-IPCEAVTVYATERYEAPELLVGTQY 179

Db 121 RVLHRLDKPQNLIRKKGELK-ADFGARAGIPVITYSHEVVT-VMWRAPDLVNGSKRYK 180

QY 180 SSVDIWAIGCVFAELLTGPLMPGKSDVDQLYLI-RTVETGFHRHVDQAGLEJLTSSDPP 239

Db 181 TPID:WSALCIFAEWASGRPL:FFGSGTSDQLFRIFKLJGP- - - - -NESSWP 227

QY 24C AVASGAGITGKLI:PHOS:F- - - - -KSNGFHGIS:PEPEDMETLEKESDVHPVALN 293

Db 22B SITE- - - - -LPKYTKDFPVH:FAHGLSSIVHG- - - - -LDEK- - - - -GLN 260

QY 294 FMKGGLKNNPDRRLTCSQLLESSYDSEFQ 322

Db 261 LLSXMLEQYDPNQRITAAALAKHPYFDGLE 289

RESULT 7

CDK2_HUMAN

ID _CDK2_HUMAN STANDARD; PRT; 298 AA.

AC P24941;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Cell division protein kinase 2 (EC 2.7.1.-) (p33 protein kinase).

GN CDK2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CX NCBI_TaxID=9606;

RN [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=911330891; PubMed=1714386;

RX Elledge S.J., Spottswood M.R.;

RA "A new human p34 protein kinase, CDK2, identified by complementation

RT of a cdc28 mutation in *Saccharomyces cerevisiae*, is a homolog of

RT Xenopus Egi.;"

RL EMBO J. 10:2653-2659(1991).

RN [2]

RN SEQUENCE FROM N.A.

RP MEDLINE=91367262; PubMed=1653904;

RX Tsai L.-H., Harlow E., Meyerson M.;

RA "Isolation of the human cdk2 gene that encodes the cyclin A- and

RT adenovirus E1A-associated p33 kinase.;"

RL Nature 353:174-177(1991).

RN [3]

RN SEQUENCE FROM N.A.

RP MEDLINE=92020980; PubMed=1717994;

RX Ninomiya-Tsuji J., Nomoto S., Yasuda H., Reed S.I., Matsumoto K.;

RA "Cloning of a human cDNA encoding a CDC2-related kinase by

RT complementation of a budding yeast cdc28 mutation.;"

RL Proc. Natl. Acad. Sci. U.S.A. 88:9006-9010(1991).

RN [4]

RN SEQUENCE FROM N.A.

RP Kieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,

RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,

RA Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;

Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RN [5]

RN SEQUENCE FROM N.A.

RP TISSUE=Placenta;

RC MEDLINE=2388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan T., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Marzly D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kesteven M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra V.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.;"

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [6]

RN PHOSPHORYLATION SITES.

RP MEDLINE=93010995; PubMed=1396589;

RX Gu Y., Rosenblatt J., O'Morgan D.O.;

RA "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160

RT and Tyr15.;"

RL EMBO J. 11:3995-4005(1992).

RN [7]

RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).

RP MEDLINE=93288132; PubMed=8510751;

RA de Bondt H.L., Rosenblatt J., Jancarik J., Jones H.D.,

RA Morgan D.C., Kim S.-H.;

RT "Crystal structure of cyclin-dependent kinase 2.;"

RL Nature 363:595-602(1993).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CYCLIN A.
 RX MEDLINE=95356811; PubMed=7630397;
 RA Jeffrey P.D., Russo A.A., Polyak K., Gibbs E., Hurwitz J.,
 RA Massague J., Pavletich N.P.;
 RT "Mechanism of CDK activation revealed by the structure of a
 RT cyclinA-CDK2 complex.";
 RL Nature 376:313-320(1995).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
 RX MEDLINE=96181476; PubMed=8610113;
 RA De Azevedo W.F., Jr., Muleer-Diekmann H.-J., Schulze-Gahmen U.,
 RA Worland P.J., Sausville E., Kim S.-H.;
 RT "Structural basis for specificity and potency of a flavonoid
 RT inhibitor of human CDK2, a cell cycle kinase.";
 RJ Proc. Natl. Acad. Sci. U.S.A. 93:2735-2740(1996).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CG2A AND K-P1.
 RX MEDLINE=96330318; PubMed=8684460;
 RA Russo A.A., Jeffrey P.D., Patten A.K., Massague J., Pavletich N.P.;
 RT "Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor
 RT bound to the cyclin A-Cdk2 complex.";
 RL Nature 382:325-331(1996).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
 RX MEDLINE=96331126; PubMed=8756328;
 RA Russo A.A., Jeffrey P.D., Pavletich N.P.;
 RT "Structural basis of cyclin-dependent kinase activation by
 RT phosphorylation.";
 RJ Nat. Struct. Biol. 3:696-700(1996).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=97075215; PubMed=8917641;
 RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;
 RT "High-resolution crystal structures of human cyclin-dependent kinase
 RT 2 with and without ATP: bound waters and natural ligand as guides for
 RT inhibitor design.";
 RL J. Med. Chem. 39:4540-4546(1996).
 RN [13]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=97475219; PubMed=9334743;
 RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,
 RA Erdicort J.A.;
 RT "Protein kinase inhibition by staurosporine revealed in details of
 RT the molecular interaction with CDK2.";
 RL Nat. Struct. Biol. 4:796-801(1997).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CKS1.
 RX MEDLINE=96182647; PubMed=8601310;
 RA Bourne Y., Watson M.H., Hickey M.J., Holmes W., Rocque W., Reed S.I.,
 RA Tainer J.A.;
 RT "Crystal structure and mutational analysis of the human CDK2 kinase
 RT complex with cell cycle-regulatory protein Cks1.";
 RL Cell 84:863-874(1996).
 RN [15]
 RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
 RX MEDLINE=98342369; PubMed=9677190;
 RA Gray N.S., Wodicka L., Thunnissen A.-M.W.H., Norman T.C., Kwok S.,
 RA Espinoza F.H., Morgan D.O., Barnes G., Leclerc S., Meijer L.,
 RA Kim S.H., Lockhart D.J., Schultz P.G.;
 RT "Expliciting chemical libraries, structure, and genomics in the search
 RT for kinase inhibitors.";
 RL Science 281:533-538(1998).
 CC - FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
 CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
 CC DURING S PHASE AND G2.
 CC - ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
 CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-163 ACTIVATES IT.
 CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC3/CDKX SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X61622; CAA43807.1; -
 CC EMBL; X62071; CAA43985.1; -
 CC EMBL; M68520; AAA35667.1; -
 CC EMBL; AF512553; AAM34794.1; -
 CC EMBL; BC003065; AAH03065.1; -
 CC PIR; A41227; A41227.
 CC PDB; 1FIN; 27-JAN-97.
 CC PDB; 1HCK; 07-DEC-96.
 CC PDB; 1HCL; 07-DEC-96.
 CC PDB; 1AQ1; 12-NOV-97.
 CC PDB; 1JST; 11-JAN-97.
 CC PDB; 1JSU; 29-JUL-97.
 CC PDB; 1BUH; 09-SEP-98.
 CC PDB; 1B38; 23-DEC-98.
 CC PDB; 1B39; 23-DEC-98.
 CC PDB; 1CKP; 13-JAN-99.
 CC PDB; 1D18; 28-JAN-03.
 CC PDB; 1DM2; 31-MAY-00.
 CC PDB; 1E1V; 10-MAY-01.
 CC PDB; 1E1X; 10-MAY-01.
 CC PDB; 1FSQ; 27-DEC-00.
 CC PDB; 1FQ1; 09-MAY-01.
 CC PDB; 1FVT; 17-JAN-01.
 CC PDB; 1FVU; 17-JAN-01.
 CC PDB; 1GSS; 14-NOV-01.
 CC PDB; 1G1H; 06-FEB-02.
 CC PDB; 1G1I; 06-NOV-02.
 CC PDB; 1G1J; 06-FEB-02.
 CC PDB; 1GY3; 06-JUN-02.
 CC PDB; 1H1P; 19-SEP-02.
 CC PDB; 1H1Q; 19-SEP-02.
 CC PDB; 1H1R; 19-SEP-02.
 CC PDB; 1H1S; 19-SEP-02.
 CC PDB; 1H24; 01-FEB-03.
 CC PDB; 1H25; 01-FEB-03.
 CC PDB; 1H26; 01-FEB-03.
 CC PDB; 1H28; 01-FEB-03.
 CC PDB; 1USV; 29-AUG-01.
 CC PDB; 1JVP; 21-DEC-01.
 CC PDB; 1KE5; 14-MAY-02.
 CC PDB; 1KE6; 14-MAY-02.
 CC PDB; 1KE7; 14-MAY-02.
 CC PDB; 1KE8; 14-MAY-02.
 CC PDB; 1KE9; 14-MAY-02.
 CC -----
 CC Query Match 30.4%; Score 554; DB 1; Length 298;
 CC Best Local Similarity 38.4%; Pred. No. 2.9e-32;
 CC Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
 CC
 CC QY 1 MEKYEKLAKTGTGSGYGVVFKCRNKTSGOVVAVKKEVEDDPVVKIALREIRMLKQLKH 60
 CC DB 1 MENQKVEKIGEGYGVVYKARNKLTGEVVALKIRLCTETEGVPSTAIRISLLKELNH 60
 CC QY 61 PNLVNLVEVFRKRKRMHLVEFYCDHTLLNLELRNP NGVADGVKSVLWQTQALNFCFI 119
 CC DB 61 PNVKLLDVIHTENKLYLVFEFLHQDLKPKMDASALTGIPLPKLSYFOLLQGLAFCHS 120
 CC QY 120 HNCIHRDLKPNILITKQIIKICDPCFAQIL-IPGDATDYVATWYRAPPELLVGDTOY 178
 CC DB 120 HRVLHRDLKPNLINTEGAIKLADFLARAFGVPRVITYTHEVVTWLYRAPELLGCKYY 180
 CC QY 179 GSSVDVIAWGCVFAELLTGQPLWPKGSDVQLYLIRTL----- 217
 CC DB 181 STAVDINSLGCIFAEMVTRRALPFGDSEIDQLFRIFRTLTGTPDEVVWVGVTSMPDKPSF 240
 CC QY 218 ---VETGRHV-----DOAGLELLTSS---DPPAVASQASGITGKLIPIRHSIFKSNFFH 267

```
Db 241 PKWARQDFSKVVPPLDEGRSLLSQWLYDPNKRISAKAAL-----AHFFQ 267
QY 268 GISIEP 274
Db 288 DVTKPE 294

RESULT 8
CC2: CRYSA
ID CC21 CRYSA STANDARD; PRT; 294 AA.
AC P29618;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Cell division control protein 2 homolog 1 (EC 2.7.1.-).
GN CDC2-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Cryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RX MEDLINE=92293101; PubMed=1376401;
RA Hashimoto J., Hirabayashi T., Hayano Y., Hata S., Ohashi Y.,
RA Suzuki I., Utsugi T., Toh-E A., Kikuchi Y.;
RT "Isolation and characterization of cDNA clones encoding cdc2
RT homologues from Oryza sativa: a functional homologue and cognate
RT variants.";
RL Mol. Gen. Genet. 233:10-16(1992).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X60374; CAA42922.1; -
CC PIR; S22440; S22440.
CC HSSP; P24941; IHCL.
CC
CC Gramine; P29618;
CC
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; Phosphorylation.
CC DOMAIN 4 287
CC NP_BIND 10 18 ATP (BY SIMILARITY).
CC BINDING 33 33 ATP (BY SIMILARITY).
CC ACT_SITE 127 127 BY SIMILARITY.
CC MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 16 16 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC SEQUENCE 294 AA; 34071 MW; 51322D93A6F4C131 CRC64;
Query Match 30.3%; Score 551.5; DB 1; Length 294;
```

```
Best Local Similarity 34.8%; Pred. No. 4.3e-32;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQWAVAKVFSEDDPVYKXIALREIRMLKQLKH 60
Db 1 MEQYEEKIKGGTYGVYRARDKVNTETIAKKIRLEQDEGVFPSTAIRISLKEKHH 60
QY 61 PNLVNLIEVFRRKRKHVLFVEYCDHTLLN-----ELERNENGVAQGVKSVLMQTLQAL 114
Db 61 GNIVRLHDVHSEKRIYLVFEYLDJLKKFNDSCPEFAKNPT-----LIKSYLVQILRGV 115
QY 115 NFCHIHNCIHRDIPEN::::TKQ-GIKICDGFQAQL-IPCDAYTYVATRWRAPELL 172
Db 115 AYCHGRVLRDLKPNLIDRRTNALKLADFLARAFGIPVRTTFHEVTLWYRAPEIL 175
QY 173 VGDITGVSSVDIWAIGCVFAELLTQPLWPGKSDVDQLYLIIRTLVETGFRHVSCAGLEL 232
Db 173 LGSROYSTPDMMSVGCIFAEMVQKPLFGDSIDELFKIFRVLGTP----- 223
QY 233 LTSSDPPAVASOSAGITGKL:PRHQSPKSGNPFHGHSIPEDMETLEEKFSOVHVAL 292
Db 233 -NEQSWPGVSS-----LPDYKSAFPK-----WQAQDLATI---VPTLDPAGL 261
QY 293 NFMKGLKMNPPDRLTCSQLLESSYFDSFOEAQ 325
Db 262 DLSQMLRVEPNKRITARQALEHEYFKDLEWVQ 294

RESULT 9
CC2A ARATH STANDARD; PRT; 294 AA.
ID CC2A ARATH STANDARD; PRT; 294 AA.
AC P24100; Q9M307;
DT 01-MAR-1992 (Rel. 21, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Cell division control protein 2 homolog A (EC 2.7.1.-).
GN CDC2A OR CDC2 OR AT3G48750 OR T21J18.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93005715; PubMed=1840925;
RA Ferreira P.C.G., Hemery A.S., Villarroel R., van Montagu M., Inze D.;
RA "The Arabidopsis functional homologue of the p34cdc2 protein kinase.";
RL Plant Cell 3:531-540(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=92039027; PubMed=1937013;
RA Hirayama T., Imaizumi Y., Anai T., Matsui M., Oka A.;
RT "Identification of two cell-cycle-controlling cdc2 gene homologs in
RT Arabidopsis thaliana.";
RL FEBS Lett. 304:73-77(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316202; PubMed=1618302;
RA Imaizumi Y., Hirayama T., Endoh H., Oka A.;
RT "Exon-intron organization of the Arabidopsis thaliana protein kinase
RT genes CDC2a and CDC2b.";
RL FEBS Lett. 304:73-77(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339744; PubMed=1634002;
RA Inze D., Ferreira P.C.G., Hemery A.S., van Montagu M.;
RT "Control of cell division in plants.";
RL Biochem. Soc. Trans. 20:80-84(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
```


or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL; AJ223949; CAA11680.1; -.
CC HSSP; P24941; IAQ1.
CC InterPro; IPR000719; Prot Kinase.
CC InterPro; IPR002290; Ser Thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
CC Cell division; Mitosis; Phosphorylation; Alternative splicing.
CC DOMAIN 4 296 PROTEIN_KINASE.
CC NP_BIND 10 18 ATP (BY SIMILARITY).
CC BINDING 33 33 ATP (BY SIMILARITY).
CC ACT_SITE 127 127 BY SIMILARITY.
CC MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC SEQUENCE 298 AA; 33873 MW; 3B580D8C2460D8E CRC64;

Query Match 30.2%; Score 550; DB 1; Length 298;
Best Local Similarity 38.1%; Pred. No. 5,6e-32;
Matches 117; Conservative 55; Mismatches 89; Indels 46; Gaps 6;

Qy 1 MEKYEKAKTGGSGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKOLKH 60
Db 1 MENFKVEKIGEGTYGVVYKAKNLTGEVVALKIRLDTETEGVPS*AIRESISLLKELNH 60

Qy 61 PNLVNLIEVFRKRRKMHVFEYCHTLLNELRNP-NCVAGVIVKSVLWOTLQALNECHI 119
Db 61 PNVKLLSDVIHTENKLYLVFFLHQDLKKFMDASAVTGIPULIKSVLFLQLQGFCHS 120

Qy 120 HNCIHRDIPKNILITKGIKICDFGFAQL-IPGDAYTDYVATRYRAPELLVGDTOY 178
Db 121 HRVLRHOLKPNOLLINAGSIIKADFGARAFGVPRYITHEVVTLYWRAPELLGCKY 180

Qy 179 GSSVDIWAIGCVFAELLTGQPLPKGSVDVQYLIIRTL----- 217
Db 181 STAVDINSIGCFIEMVTRRALFGDSEIDQLFRIFRTLGTDPVWPGVTSMPPYKPSF 240

Qy 218 ---VETGRHV-----DQAGLELLTSS---DPPAVASQAGITGKLIIPRHQSIFKNSGFFH 267
Db 241 PKWARQDSKVPPLCEDGRLSSQMLHYDNPKRISAKAAL-----ARPFQ 267

Qy 268 GISPEP 274
Db 268 DVTKPVP 294

```

RESULT 11

```

CCDC2_ORYJA STANDARD; PRT; 303 AA.
ID CDG2_ORYJA
AC Q9DGA2; Q9DGA1.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein
DE kinase) (Cyclin-dependent kinase 1) (CDK1).
GN CDK2.
OS Oryzias javanicus (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=123683;
RN [1]_TaxID=123683;
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yamashita M., Mita K.;
RT "cDNA cloning of Cdc2 and cyclin B in medaka species.";

```

```

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL
CC CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND
CC MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE
CC REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC
CC COMPONENT OF MPF (BY SIMILARITY).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN
CC NATURE COCYTES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcel/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB050461; BAB17219.1; -.
CC EMBL; AB050462; BAB17220.1; -.
CC HSSP; Q00534; IB18.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
CC DOMAIN 4 287 PROTEIN_KINASE.
CC NP_BIND 10 18 ATP (BY SIMILARITY).
CC BINDING 33 33 ATP (BY SIMILARITY).
CC ACT_SITE 128 128 BY SIMILARITY.
CC MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC VARIANT 49 49 V -> I.
CC VARIANT 245 245 M -> K.
CC VARIANT 269 269 I -> T.
CC VARIANT 276 276 I -> V.
CC VARIANT 299 299 C -> S.
CC SEQUENCE 303 AA; 34691 MW; 439B8B012BE28D9C CRC64;

```

Query Match 30.2%; Score 549.5; DB 1; Length 303;

Best Local Similarity 35.7%; Pred. No. 6.2e-32;

Matches 120; Conservative 71; Mismatches 94; Indels 51; Gaps 9;

```

Qy 1 MEKYEKAKTGGSGYGVVFKCRNKTSGQVAVKVF-VESEDDPVVKKIALREIRMLKOLK 59
Db 1 MEDYVVKIEKIGESTYGVVYKGRKSTSGQVAVKVKIRLESEEG-VPESTAVREVLQLQELK 59

Qy 60 HPNLVNLIEVFRKRRKMHVFEYCHTLLNELRNPNG--VADGVIVKSVLWOTLQALNFC 117
Db 60 HPNVRLDVLVMOESRLYLIFELSLMDLKKYLDSPSGQYMDPMLVKSYLQILEGIYFC 119

Qy 118 HINCCHRRDIKPNILITKGIKICDFGFAQL-IPGDAYTDYVATRYRAPELLVGD 176
Db 120 HRRVLRHRLKPNOLLINAGSIIKADFGARAFGVPRYITHEVVTLYWRAPELLGSP 179

Qy 177 QYGSVVDIWAIGCVFAELLTGQPLPKGSVDVQYLIIRTLVETGRHVDPQAGLELLTSS 236
Db 180 RYSTPVDVWSTGTIFAEATKPKPLFGDSEIDQLFRIFRTL-----G 221

Qy 237 DP-----PAVASQAGITGKLIIPRHQSIFKNSGFFHGISIPEDPME-TLEKFSQVHPV 290
Db 222 TPNDVWDPVES-----LPDYKNTF-----PKWMEGSLSSWVKLNDKN 259

```

QY 291 ALNFMKGLKLVNPDRLTCSQLLESSYFDSFOEQAQI 326
 Db 260 GLDLLAKMLJYNPKRISAREATHPVFDLAKSTL 295

RESULT 12

CDK2 RAT
 ID CDK2 RAT STANDARD; PRT; 298 AA.
 AC Q63699; Q09136;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cell division protein kinase 2 (EC 2.7.1.1).
 GN GN
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN RATTUS NORVEGICUS (RAT).
 RS SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=95166553; PubMed=7862443;
 RA Kotani S., Endo T., Kitagawa M., Higashi H., Onaya T.,
 RT "A variant form of cyclin-dependent kinase 2 (CDK2) in a malignantly
 transformed rat thyroid (FRTL-4) cell line."
 RJ Oncogene 10:663-669(1995).
 RL (2)
 RN SEQUENCE OF 19-124 FROM N.A.
 RP MEELINE=96113578; PubMed=8673024;
 RX Hosokawa Y., Yang M., Kaneko S., Tanaka M., Nakashima K.,
 RT "Synergistic gene expressions of cyclin E, cdk2, cdk5 and E2F-1
 during the prolactin-induced G1/S transition in rat Nb2 pre-T
 lymphoma cells."
 RL Biochem. Mol. Biol. Int. 37:393-399(1995).
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
 CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
 CC DURING S PHASE AND G2.
 CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
 CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
 CC SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=CDK2-alpha;
 CC IsoId=Q63699-1; Sequence=Displayed;
 CC Name=CDK2-beta;
 CC IsoId=Q63699-2; Sequence=Not described;
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC2/CDKX SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D28753; BAA05947.1; -;
 CC EMBL; D63162; BAA09638.1; -;
 CC HSP; P24941; IAO1
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMC0220; S_TKc; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
 KW Cell division; Mitosis; Phosphorylation; Alternative splicing.
 FT DOMAIN 4 286 PROTEIN KINASE.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).

FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 127 127 BY SIMILARITY.
 FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
 FT CONFLICT 79 79 V -> C (IN REF. 2).
 FT CONFLICT 99 99 L -> I (IN REF. 2).
 FT CONFLICT 124 124 L -> C (IN REF. 2).
 SQ SEQUENCE 298 AA; 33887 MW; C8CB3ADCE9B97F88 CRC64;
 Query Match 30.1%; Score 548; DB 1; Length 298;
 Best Local Similarity 38.1%; Pred. No. 7.8e-32;
 Matches 17; Conservative 55; Mismatches 89; Indels 46; Gaps 6;
 QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGOVAVKVFSEDDPVVKIALREIRMLKOLKCH 60
 Db 1 MENQKVEKIEGEGYGVVYKAKNKLTEGVKALRILDTETEGVPSTAIRISILKELNH 60
 QY 61 PNLVNLIEVFRKRKRMHLVFYCDHTLLNELRNP-NGVADGVIKSVLMQC'LQALNFCHI 119
 Db 61 PNIVKLLDVIHTENKLYLVFEFLHQDLKFKMDASALTGLPLIKSYDLFQLLOGLAFCHS 120
 QY 120 ENCIHRDIKPNILITKGGIIKICDPGPAQIL-IPGDAYTDYVATRVAPPELLVGDQY 178
 Db 121 ERVLRDRLKPNLLINAEKSLADFGUAFGVPRVITYTHEVTLWYRAPEILLGCKYY 180
 QY 179 GSSVDIWAIGCVFAELTQPLWFGKSDVDQLYLIIRTL----- 217
 Db 181 STAVDIWSLGCIFAEWVTTRALPGDSEIDQLFRITLGTDPDEVWMPGVTSMPDYKPSF 240
 QY 218 ---VETGPRHV-----DQAGLELLTSS---DPPAVASQAGITGKLIPHQIFKSNGRFH 267
 Db 241 PKWAFQDFSKVVPPLDEDSRLLSQMLHYDPNKRISAKAAL-----AHPEFFQ 287
 QY 288 GISPEP 274
 Db 288 DVTRVP 294
 RESULT 13
 CDK2 MESAU
 ID CDK2 MESAU STANDARD; PRT; 298 AA.
 AC P49363;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cell division protein kinase 2 (EC 2.7.1.1).
 GN CDK2
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [-];
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94107363; PubMed=8280171;
 RA Noguchi E., Sekiguchi T., Yamashita K., Nishimoto T.;
 RT "Molecular cloning and identification of two types of hamster cyclin-
 RT dependent kinases: cdk2 and cdk2L."
 RL Biochem. Biophys. Res. Commun. 197:1524-1529(1993).
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
 CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
 CC DURING S PHASE AND G2.
 CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
 CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC2/CDKX SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

-----
DR EMBL: D17350; BAA01465.1; --
DR HSSP: P24941; 1AQ1
DR InterPro: IPR000719; Prot_kinase
DR InterPro: IPR002290; Ser_Thr_Pkinase
DR Pfam: PF00069; Pkinase; 1
DR ProDom: PD000001; Prot_kinase; 1
DR SMART: SM00220; S_TKc; 1
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 286 PROTEIN_KINASE
FT NP_BIND 10 18 ATP (BY SIMILARITY);
FT BINDING 33 33 ATP (BY SIMILARITY);
FT ACT_SITE 127 127 BY SIMILARITY;
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY);
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY);
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY);
SQ SEQUENCE 298 AA; 33839 MW; D11C036664C47CE CRC64;

Query Match 30.0%; Score 546; DB 1; Length 298;
Best Local Similarity 38.1%; Pred. No. 1.1e-31;
Matches 117; Conservative 54; Mismatches 90; Indels 46; Gaps 6;

Qy 1 MEKYEKLAKTGEQYGVVFKRNKTSGQWAVKVFSEDDPVVKALREIMLKOLKH 60
Db 1 MENFQVEKIGEGYGVYKAKNLTGEVVALKIRLDTTEGVSTAIRISILKELNH 60
Cy 61 PNLVNLIEVFRKPKHMLFEYCDHTLLNELRNP-NGVADGVKISVLMTQLOALNFCI 119
Dz 61 PNIYKLLDVHTENKLYLVPELLHQDLKKFMDASAVTGIP-PLIKSYLFQLLGLAFCHS 120
Qy 120 HNCIHRDIKENILITKGIKICDGFQAQIL-TPGEAYTDYVATRYRAPELLVGDTQY 178
Db 121 HRVLHRLKPNQLNLIAEGSKLADPGLARAFGVFVRYTITHEVVTLYRAPELLGGKYI 180
Qy 179 GSSVDIWAIGCVFAELITGQPLWPKSDVDQYLIIRTL----- 217
Db 181 STAVDINSLGCIPEWYTRRALFGDSEIQGLFRIPLTGTPDEWNPVTS:PDYKFSF 240
Qy 218 ---VETGFRHY-----DQAGLELLTSS---DPPAVASOSAGITGKLI PRHOS:FKNSGFFH 267
Db 241 PKWARQDFSKYVPLDDEGRSLLSQMLHYDPNKRISAKAL-----AHPFFQ 287
Qy 268 GISIPEP 274
Db 288 DVTKPVP 294

RESULT 14
CJC2 HUMAN
IC -CDC2 HUMAN STANDARD; PRT; 297 AA.
AC PC6493; O60764;
DT 01-JAN-1998 (Rel. 06, Created)
DT 01-JAN-1998 (Rel. 06, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein
DE kinase) (Cyclin-dependent kinase 2) (CDK2).
GN CDC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87201915; PubMed=3553962;
RA Lee M.G., Nurse P.;
RT "Complementation used to clone a human homologue of the fission yeast

```

```

RT cell cycle control gene cdc2.";
RL Nature 327:31-35(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT CDC2DELTAT.
RC TISSUE=Breast cancer;
RX MEDLINE=98175539; PubMed=9515786;
RA Ohta T., Okamoto K., Isohashi F., Shibata K., Fukuda M., Yamaguchi S.,
RA Xiong Y.;
RT "A loop deletion of CDC2 from breast cancer tissues eliminates binding
RT to cyclin B1 and cyclin-dependent kinase inhibitor p21.";
RL Cancer Res. 58:1095-1098(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Bieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Wiatr L.A., Nickerson D.A.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22382857; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP PHOSPHORYLATION, AND ASSOCIATION WITH P13.
RX MEDLINE=88253421; PubMed=3289755;
RA Braetta G., Beach D.;
RT "Activation of cdc2 protein kinase during mitosis in human cells:
RT cell cycle-dependent phosphorylation and subunit rearrangement.";
RL Cell 54:17-26(1988).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE
CC AND MITOSIS. P34 IS A COMPONENT OF THE KINASE COMPLEX THAT
CC PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA
CC POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT.
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY
CC SUBUNIT AND WITH A CYCLIN.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- POLYMERPHISM: Isoform CDC2delta is found in breast cancer
CC tissues. It is unable to complex with cyclin B1 and also fails to
CC bind to the CDK inhibitor p21.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR EMBL: X05360; CAA28963.1; -.

```

DR EMBL; Y0272; CAA68376.1; -.
DR EMBL; D88357; BAA26001.1; -.
DR EMBL; AF12854; AAM34793.1; -.
DR EMBL; BC014563; AAM14563.1; -.
DR P.R; A29519; A29539.
DR PDB; 1LC9; 17-APR-02.
DR SWISS-2DPAGE; P06493; HUMAN.
DR Genew; HNCI:1722; CDC2.
DR Gk; P06493; -.
DR MIM; 116940; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Cell division; Mitosis; Nuclear protein; Phosphorylation;
KW Polymorphism; 3D-structure.
FT DOMAIN 4 287 PROTEIN KINASE.
FT NP_BIND 10 16 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT VARIANT 107 163 MISSING (IN CDC2delta7).
FT FT
FT SEQUENCE 297 AA; 34095 MW; 942D79448EFE490A CRC64;

Query Match 29.9%; Score 545; DB 1; Length 297;
Best Local Similarity 40.2%; Pred. No. 1.3e-31;
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGOVAVKVFVESEDDPVVKKIALRIRMKQLKH 59
DB 1 MEDYTKIEKIGEGYGVVYKRNKTKTQGVVAMKRLSEEEGVFSTAIRSLSJLKEJR 59

QY 60 HPNLVNLIEVFRKRKRKMLHVFYECCHTLNLELERNFNG--VADGVKSVLWOTLQALNFC 117
DB 60 HPNIVSLQDVLMQDSRLYLFEFLSMGLKYLDS--PGQVYDSSANKSVLYQLQGVIVC 119

QY 118 HINCVRDRIKPNILITKQGIKICDFGPAQIL-IPGDHYTDYVATRWYRAPELLVGT 176
DB 120 HSRVLHRLDKPQNLG--DDKGTIKLADFLARAFGPIRVYTHEVVTWLYRSPEVLGSA 179

QY 177 QYGVSSVDIWAIGCVFAELLTGQPLWPKGSDVDQLYLIIRT----- 217
DB 180 RYSTPVDIWSIGTIFAEALATKPLFHGDSIDQLFRIFALGTPTNNEVPEVESJQYQXN 239

QY 218 -----VETGFRHVDQAGLELLTSS-----DPVAVASQASAGITGLIPRH 256
DB 240 TFPKWKPGSLASHVQNLDENGDLSKMLYDP-----AKRISGKVALNH 284

RESULT 15
CDK2_CARAU STANDARD; PRT; 298 AA.
>D CDK2_CARAU
AC P43450;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 13-JUL-1999 (Rel. 38, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.1.-).
GN CDK2.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Cocyte;
RX MEDLINE=92331802; PubMed=139336;
RA Hirai T., Yamashita Y., Yoshikuni M., Tokumoto T., Kajiura H.,
RA Sakai N., Nagahama Y.;
RT "Isolation and characterization of goldfish cdk2, a cognate variant
RT of the cell cycle regulator cdc2.";
RL Dev. Biol. 152:113-120(1992).
CC !- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC !- INTERACTS WITH CYCLINS A, D, OR E.
CC !- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
CC SIMILARITY).
CC !- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; S40289; AAB22550.1; -.
CC FIR; A44878; A44878.
CC HSSP; P24941; IFIN.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_Thr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 286 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT SEQUENCE 298 AA; 33998 MW; 50A93A63A235F9A9 CRC64;

Query Match 29.9%; Score 545; DB 1; Length 298;
Best Local Similarity 46.6%; Pred. No. 1.3e-31;
Matches 102; Conservative 48; Mismatches 67; Indels 2; Gaps 2;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGOVAVKVFVESEDDPVVKKIALRIRMKQLKH 60
DB 1 MESFOKVEKIGEGYGVVYKRNKTKTQGVVAMKRLSEEEGVFSTAIRSLSJLKEJNH 60

QY 61 PNLVNLIEVFRKRKRKMLHVFYECCHTLNLELERNP-NGVADGVKSVLWOTLQALNFC 119
DB 61 PNIUKLHVDIHTENKLYLVFEFLHQLDKRFDSDSTVTGISLPLVKSYLEQLQGLAFCHS 120

QY 120 HNCVRDRIKPNILITKQGIKICDFGPAQIL-IPGDHYTDYVATRWYRAPELLVGTQY 178
DB 121 HRLVHRLDKPQNLG--DDKGTIKLADFLARAFGPIRVYTHEVVTWLYRAPELLGCKYY 180

QY 179 GSSVDIWAIGCVFAELLTGQPLWPKGSDVDQLYLIIRT 217
DB 181 STAVDIMSGLGICFAEMITRKALFFGDSIDQLFRIFRL 219

Search completed: November 13, 2003, 14:19:19
Job time : 11.4834 secs

```

GenCore version: 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:16:06 ; Search time 33.0227 Seconds
(without alignments)
2711.600 Million cell updates/sec

Title: US-09-671-050-10

Perfect score: 1820
Sequence: 1 MEKYSLAKTGGSGYGVFK.....RKARREGRRRRQGVLPKLS 347

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 25805264 residues

Total number of hits satisfying chosen parameters: 83525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.23.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*
15: sp.virus.*
16: sp.bacteriap.*
17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1180	64.8	352	11 Q8CEQ0	Q8ceq0 mus musculus
2	1001.5	55.0	353	5 Q912H1	Q912h1 caenorhabdi
3	1001.5	55.0	392	5 Q9VMN3	Q9vmn3 drosophila
4	941.5	51.7	566	6 Q9TTK0	Q9ttk0 cryptolagus
5	915.5	50.3	493	4 Q92772	Q92772 homo sapien
6	904.5	49.7	564	11 Q9GY12	Q9gy12 mus musculu
7	904.5	49.7	568	11 Q9QUK0	Q9qk0 mus musculu
8	902	49.6	329	11 Q9QY11	Q9qy11 mus musculu
9	758.5	41.7	455	4 Q9P114	Q9p114 homo sapien
10	758.5	41.7	592	4 Q81VW4	Q81vw4 homo sapien
11	755	41.5	657	5 Q9BMG2	Q9bmg2 trypanosoma
12	744	40.9	353	11 Q8K134	Q8kl34 mus musculu
13	743.5	40.9	1106	5 Q9GRT9	Q9grt9 leishmania
14	741.5	40.7	457	11 Q8BKR2	Q8bkr2 mus musculu
15	738.5	40.6	595	11 Q8BL49	Q8bl49 mus musculu
16	737.5	40.5	457	11 Q9JM02	Q9jm02 rattus norv

17	737.5	40.5	505	11 Q9JMC1	Q9jmc1 rattus norv
18	733.5	40.3	595	11 Q8BLF2	Q8blf2 mus musculu
19	694.5	38.2	783	11 Q8BWI8	Q8bwi8 mus musculu
20	694.5	38.2	831	4 Q81YC7	Q81yc7 homo sapien
21	682	37.5	993	4 Q8WXQ5	Q8wxq5 homo sapien
22	676.5	37.2	578	11 Q8BVE0	Q8bvec mus musculu
23	618	34.0	1104	13 Q9W6R6	Q9w6r6 fugu rubrip
24	615.5	33.8	997	4 Q9UJL6	Q9ujl6 homo sapien
25	565.5	31.1	294	10 Q9ZRI1	Q9zrl1 triticum ae
26	551	30.3	288	5 Q96821	Q96821 plasmodium
27	551	30.3	288	5 Q9XZD6	Q9xzd6 plasmodium
28	545.5	30.0	294	10 Q82666	Q82666 brassica na
29	542.5	29.8	300	5 Q15890	Q15890 toxoplasma
30	539.5	29.6	294	10 Q40790	Q40790 pinus conto
31	539.5	29.6	297	10 Q8RX68	Q8rx68 arabidopsis
32	539.5	29.6	300	5 Q17066	Q17066 asterina pe
33	539.5	29.6	300	5 Q44000	Q44000 toxoplasma
34	539	29.6	298	5 Q27032	Q27032 theileria p
35	538.5	29.6	294	10 Q40789	Q40789 petroselinu
36	538.5	29.6	294	10 Q82135	Q82135 pisum sativ
37	538.5	29.6	300	3 Q13379	Q13379 pneumocysti
38	538.5	29.6	300	3 Q13380	Q13380 pneumocysti
39	537	29.5	298	5 Q96820	Q96820 plasmodium
40	535.5	29.4	294	10 Q43361	Q43361 picea abies
41	535.5	29.4	294	10 Q91556	Q91556 sesbania ro
42	534.5	29.4	294	10 Q9FUR4	Q9fur4 nicotiana t
43	534.5	29.4	294	10 Q40484	Q40484 nicotiana t
44	534	29.3	298	5 Q26671	Q26671 theileria a
45	533.5	29.3	294	10 Q816T8	Q816t8 daucus caro

ALIGNMENTS

RESULT 1

Q8CEQ0 PRELIMINARY; PRT; 352 AA.
ID Q8CEQ0
AC Q8CEQ0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Cyclin-dependent kinase-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK016781; BAC25497.1; --
SQ SEQUENCE 352 AA; 41023 MW; 6E9F56C8080F35E0 CRC64;

Query Match 64.8%; Score 1180; DB 11; Length 352;
Best Local Similarity 62.7%; Pred. No. 6.1e-98;
Matches 222; Conservative 42; Mismatches 48; Indels 42; Gaps 4;
QY 1 MEKYEKLAKTGGSGYGVFKRNKTSGGVAVKKESEDDPVPKKIALREIRMLKQLKH 60
Db 1 MEKYEKIGKIGSGYGVFKRNRTGQIVAKRLEFEDDPVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLLNELERNPNPGVADGVKSLVMTQALNFCIH 120
Db 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLVLDRLYQGVPEPLVKNITWTQLQAVNFKKH 120
QY 121 NCIHRIKIPENILITKQGIKICDFGFAOILI-PGDATYDVATWYPAPELLVGDTOYG 179
Db 121 NCIHRIKIPENILITKQSAIKLCDFGARLLTGPDDYTYDVATWYRSPPELLVGDTOYG 180

```
QY 180 SVDYWAIGCVFAELLTQGLPFGKSDVDQLY::IRTLVETGPRHVDQAGLELLTSSPP 239
DB 181 PVDYWAIGCVFAELLTQGLPFGKSDVDQLY::IRTLVETGPRHVDQAGLELLTSSPP 218
QY 240 AVASQAGITGKLI::PRHOSIFKSNQFFHGIS::PEPEDVETLEEKFSVDVHPVALNFKGCL 293
DB 219 -----GDLIRHQVFMQVYFSVXIPQEDMETLELKFPHISYALGFLKGC 268

QY 300 KNPDRILTCSSQLLESSYFDSFQEA-----QKARNEGNNRQ-----QYLP 344
DB 269 HNDPAERLTCEQLLQHPYFDSIREYVELTRQHXKPKARKTLRQSKHUTG-QYLP 322

RESULT 2
Q9U2H1 PRELIMINARY; PRT; 353 AA.
ID Q9U2H1
AC Q9U2H1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Y42A5A.4 protein.
GN Y42A5A.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RA Stewart C.A.;
RP Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RL (2)
RM
RQ SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851915;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998)
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF032618; CAB63367.1; -.
DR HSSP; P24941; 1HCL.
DR WormPep; Y42A5A.4; CE20258.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pk_kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 353 AA; 40665 MW; 7E32260D9849CAC9 CRC64;

Query Match 55.0%; Score 1001.5; DB 5; Length 353;
Best Local Similarity 53.4%; Pred. No. 7.5e-82;
Matches 187; Conservative 59; Mismatches 63; Indels 41; Gaps 3;

QY 1 MEKYEKLTGEGSYGVWFKCRKNTSGOVAVKVFESDDPWVKALREIRMLKQLKH 60
DB 1 MKYDKSLKLGESYGVVTKCRNRTGQVAIKKFVEDDPEIKKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKRMHLVFYCDHTLLMELEERNNGVADGVKSVLQTLCAINFCH 120
DB 61 QNLVGLIEVFRRKRLHLVFLCDRTLVLHELEKNPHGVNDELKKIIVQLLEALKFCHSH 120

QY 121 NCTHRDKPENILITVQGIKICDFEFAILLPGADYDYVATRYRABELLVGDTOYGS 180
DB 121 KC1HRDVKPENIFLITNDQVKG-GDGFARINTTMYTDYVATRYRSPELLVGDVQYGP 180

QY 181 SVDIWAIGCVFAELLTQGLPFGKSDVDQLY::IRTLVETGPRHVDQAGLELLTSSOPPA 240
DB 181 PVDIWAIGCVFAELLTQGLPFGKSDVDQLY::IRTLVETGPRHVDQAGLELLTSSOPPA 217
```

```
QY 241 VASQAGITGKLI::PRHOSIFKSNQFFHGIS::PEPEDVETLEEKFSVDVHPVALNFKGCL 303
DB 218 -----GDLIRHQVFMQVYFSVXIPQEDMETLELKFPHISYALGFLKGC 268

QY 301 KNPDRILTCSSQLLESSYFDSFQEAQIKRKARNEG-----NNRRRQYVLP 345
DB 269 MSPDRFSCBELMLHGFISNW-----ILIRQDESTPGLTSKSPNVLPL 314

RESULT 3
Q9VWN3 PRELIMINARY; PRT; 392 AA.
ID Q9VWN3
AC Q9VWN3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG7236 protein.
GN CG7236.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RQ SEQUENCE FROM N.A.
RX STRAIN=SERKELEY;
RA YEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Mar K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos B.L.G.,
RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Bai-lew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cauley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Pedson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Turbin K.C., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.C., Wei M.-H., Ibegwam C.,
RA Calali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzty D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swires K., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage I., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC003611; RAFS2279.1; -.
DR HSSP; P24941; 1HCL.
DR FlyBase; FBgn0031730; CG7236.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pk_kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
```

SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 392 AA; 44962 MW; C94270077784E55D CRC64;

Query Match: 55.0%; Score 1001.5; DB 5; Length 392;
 Best Local Similarity 53.0%; Pred. No. 8.7e-82;
 Matches 184; Conservative 65; Mismatches 61; Indels 37; Gaps 3;

QY 1 MEKYEKLAKTGSGYGVFKCRNKTSQGVAVYKVFESDDPVVAKIALRIRMLKQLKH 60
 DB 1 MEKYEKLAKTGSGYGVFKCRNKTSQGVAVYKVFESDDPVVAKIALRIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKKRWHLVFEYCDHTLLNELEERNNGVADGVKSVLWQTLQALNFCIH 120
 DB 61 PNLVNLIEVFRKKRWHLVFEYCDHTLLNELEERNNGVADGVKSVLWQTLQALNFCIH 120
 QY 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYTDYVATWYRAPELLVGDYQY 179
 DB 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYTDYVATWYRAPELLVGDYQY 179
 QY 180 SSVDIWAIGCVFAELTGQPLWPGKSDVDQVLYLIRLTVEIGFRHVDQAGLELLTSSDPP 239
 DB 180 SSVDIWAIGCVFAELTGQPLWPGKSDVDQVLYLIRLTVEIGFRHVDQAGLELLTSSDPP 239
 QY 240 AVASOSAGITGKLIPRHQSIFKSNQFFHGISIPEPEMETLEEKFSVHPVALNPMKCL 299
 DB 240 AVASOSAGITGKLIPRHQSIFKSNQFFHGISIPEPEMETLEEKFSVHPVALNPMKCL 299
 QY 300 KXNPDRLTCSQLLESSYF--DSF-----OEAIK--RKARN-----EGRNRROQ 341
 DB 300 KXNPDRLTCSQLLESSYF--DSF-----OEAIK--RKARN-----EGRNRROQ 341
 QY 324 HVDPRKRPFCALLHDFQMDGFAERFSQELQMKVQKDARNISLSKKSQRKKREK 324
 DB 324 HVDPRKRPFCALLHDFQMDGFAERFSQELQMKVQKDARNISLSKKSQRKKREK 324

RESULT 5
 Q92772 PRELIMINARY; PRT; 493 AA.
 AC Q92772; 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE P56 KKIAMRE protein kinase.
 GN KKIAMRE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97152547; PubMed=9000130;
 RA Taglienti C.A., Wyck M., Davis R.J.;
 RT "Molecular cloning of the epidermal growth factor-stimulated protein
 kinase p56 KKIAMRE."
 RL Oncogene 13:2563-2574 (1996).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; U35146; AAC50918.1; ..
 DR HSSP; P24941; 1HCL.
 DR Genew; HGNC:1782; CDKL2.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 493 AA; 56018 MW; 5CC20A91CBF89EFC CRC64;

Query Match 50.3%; Score 915.5; DB 4; Length 493;
 Best Local Similarity 48.3%; Pred. No. 6.8e-74;
 Matches 172; Conservative 68; Mismatches 69; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGSGYGVFKCRNKTSQGVAVYKVFESDDPVVAKIALRIRMLKQLKH 60
 DB 1 MEKYEKLAKTGSGYGVFKCRNKTSQGVAVYKVFESDDPVVAKIALRIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKKRWHLVFEYCDHTLLNELEERNNGVADGVKSVLWQTLQALNFCIH 120
 DB 61 PNLVNLIEVFRKKRWHLVFEYCDHTLLNELEERNNGVADGVKSVLWQTLQALNFCIH 120
 QY 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYTDYVATWYRAPELLVGDYQY 179
 DB 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYTDYVATWYRAPELLVGDYQY 179

```
Db 121 NIHRDIKPNILVSGSGWKLCDGFGARTLAAPGEVTVDYVATRWYRAPELLVGVGVYK 180
QY 180 SSVDIWAIGCVFAELLTCQPLWPKGSVDQLYLIRTLVEGFPHVDQAG:ELLTSSDPP 239
Db 181 KAVDIWAIGCVLTFNFMGEPLFGSDIDQLJYHIMNCL----- 218
QY 240 AVASQAGITGLIPRHOSIFKSNFGFHGISIPEDMET--LEEKFSVHPVALNFMKGC 299
Db 219 -----GNLIPRHQELFNKQVPFAGVRLPEIKEREP--ERYPKLSEWIDLAKKCL 268
QY 300 KWPDDRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGNRRRQ 341
Db 269 HIDPDKRPFCALDHDFQMDGFAERFSQELQKIEKDARNNSLPKKSQNRKKEK 324

RESULT 6
C9QYI2 PRELIMINARY; PRT; 564 AA.
AC Q9QYI2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KXIARE-gamma.
GN CDKL2 OR KX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; and C57BL/6;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KXIARE gene: variants, dual promoters, expression and
chromosomal localization.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029066; BAA88428.1; -.
DR HSSP; P24941; 1HCL.
DR XGD; MGI:1858227; Cdkl2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PFO0069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 564 AA; 63640 MW; DACDEF630CCE82D5 CRC64;

Query Match 49.7%; Score 904.5; DB 11; Length 564;
Best Local Similarity 48.6%; Pred. No. 8,1e-73;
Matches 174; Conservative 66; Mismatches 69; Indels 45; Gaps 7;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALRIRMLKQLKE 60
Db 1 MEKYNLGLVGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALRIRMLKQLRH 60
QY 61 PNLVNLIEVFRKRKRMHLVFCYCHTLNELERNPNVGADVGVKSVLWQTLQALNFCCHI 120
Db 61 ENLVNLLEVCCKKRWYLVFVVDHTLDDKLFPNGLDYQVQKYLFIQINGIGFCHSH 120
QY 121 NCIHRIKPNILITKGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCYG 179
Db 121 NIHRDIKPNILVSGSGWKLCDGFGARTLAAPGEVTVDYVATRWYRAPELLVGVGVYK 180
QY 180 SSVDIWAIGCVFAELLTCQPLWPKGSVDQLYLIRTLVEGFPHVDQAG:ELLTSSDPP 239
Db 181 KAVDIWAIGCVLTFNFMGEPLFGSDIDQLJYHIMNCL----- 218
QY 240 AVASQAGITGLIPRHOSIFKSNFGFHGISIPEDMET--LEEKFSVHPVALNFMKGC 297
Db 219 -----GNLIPRHQELFNKQVPFAGVRLPEIKEREP--ERYPKLSEWIDLAKKCL 268
QY 297
Db 297
```

```
QY 298 CLKMPPDRRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGNRRRQ 341
Db 269 CLKMPPDRRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGNRRRQ 326

RESULT 7
C9QUK0 PRELIMINARY; PRT; 568 AA.
AC Q9QUK0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE SER/THR kinase KXIARE.
GN CDKL2 OR KX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ, and C57BL/6;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KXIARE gene: variants, dual promoters, expression and
chromosomal localization.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029073; BAA88439.1; -.
DR EMBL; AB029068; BAA88439.1; JOINED.
DR EMBL; AB023069; BAA88439.1; JOINED.
DR EMBL; AB023070; BAA88439.1; JOINED.
DR EMBL; AB029071; BAA88439.1; JOINED.
DR EMBL; AB029072; BAA88439.1; JOINED.
DR EMBL; AB029065; BAA88427.1; -.
DR HSSP; P24941; 1HCL.
DR XGD; MGI:1858227; Cdkl2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PFO0069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 568 AA; 64055 MW; A43B75E2E9EB86C4 CRC64;

Query Match 49.7%; Score 904.5; DB 11; Length 568;
Best Local Similarity 48.6%; Pred. No. 8,1e-73;
Matches 174; Conservative 66; Mismatches 69; Indels 49; Gaps 7;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALRIRMLKQLKH 60
Db 1 MEKYNLGLVGEISYGVVFKCRNKTSGQVAVKFESEDDPVVKKIALRIRMLKQLRH 60
QY 61 PNLVNLIEVFRKRKRMHLVFCYCHTLNELERNPNVGADVGVKSVLWQTLQALNFCCHI 120
Db 61 ENLVNLLEVCCKKRWYLVFVVDHTLDDKLFPNGLDYQVQKYLFIQINGIGFCHSH 120
QY 121 NCIHRIKPNILITKGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCYG 179
Db 121 NIHRDIKPNILVSGSGWKLCDGFGARTLAAPGEVTVDYVATRWYRAPELLVGVGVYK 180
QY 180 SSVDIWAIGCVFAELLTCQPLWPKGSVDQLYLIRTLVEGFPHVDQAG:ELLTSSDPP 239
Db 181 KAVDIWAIGCVLTFNFMGEPLFGSDIDQLJYHIMNCL----- 218
QY 240 AVASQAGITGLIPRHOSIFKSNFGFHGISIPEDMET--LEEKFSVHPVALNFMKGC 297
Db 219 -----GNLIPRHQELFNKQVPFAGVRLPEIKEREP--ERYPKLSEWIDLAKKCL 268
QY 297
Db 297
```

Db 269 CLHIDPKRPFCDLRLRHCHFFQMDGFAERFSQELQLKIEKDARNNSLPFKSKQNRKXK 326

RESULT 8

```
Q9QY11 PRELIMINARY: PRT: 329 AA.
AC Q9QY11;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KIAMRE-delta (Fragment).
GN CDK2 OR KKH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KIAMRE gene: variants, dual promoters, expression and
RT chromosomal localization."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB023067; BAA38429.1; -.
DR HSSP; P24941; IHCL.
DR GENE; HGNC:15483; CDKL3.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER
FT SEQUENCE 329 AA; 38023 MW; F72EB98E5A8585 CRC64;
Query Match 49.6%; Score 902; DB 11; Length 329;
Best Local Similarity 49.0%; Pred. No. 6.4e-73;
Matches 173; Conservative 63; Mismatches 73; Indels 44; Gaps 6;
QY 1 MEKYEKLANTGEGSYGVFKCRNKTSGQVAVKKVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEGLVGEYSYGVWCKRKSQGRVAIKFLSEDDDDKWKIARIRKILKQLRH 60
QY 61 PNLVNLIEVFRKRRKHVLFVEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
DB 61 ENLVNLIEVCKKKRWLVFEFVDHTLDDKLFPNGLDYQVQKYLFTQINGIGFCHSH 120
QY 121 NCIHRIKPEKNTLITKQGIKICDGFQAQILIPGDAYTDVYATWYRAPELLVGDTOYG 179
DB 121 NIHRIKPEKNTLVSGGVVKKLDCDFGFAITLAAPGDIYDVYATWYRAPELLVGDVIG 190
QY 180 SSVDIWAIGCVFAELLTGQPLWPKGSDVDQVLYIIRTLVETGFRHVDQAGLELLTSSDP 239
DB 181 KAVDIWAIGCVLWMLGQPLPFGESDIQLHHITCL----- 218
QY 240 AVASQAGITGKLIPIHQSIKSNFFHGSIPEPEDMET--LEKFSQVHFVALNFKG 297
DB 219 -----GNLIPRQHLFYKNPVFAGVLPVFKVCAEAELSESRYPKLEAVISLAKK 268
QY 298 CLKWNPDRLTCSQLSESSYF--DSF-----GDAQ-K--RKAREGRNRRQC 341
DB 269 CLHIDPKRPFCDLRLRHCHFFQMDGFAERFSQELQLKIEKDARNNSLPFKSKX 321
```

RESULT 9

```
Q9P114 PRELIMINARY: PRT: 455 AA.
AC Q9P114;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE Serine-threonine protein kinase NKIAMRE.
GN NKIAMRE
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Widmer M., Hag R., Zanke B.W.;
RC "NKIAMRE a novel kinase deleted in human leukemia."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF130372; AAF36509.1; -.
DR HSSP; P24941; IHCL.
DR GENE; HGNC:15483; CDKL3.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER
FT SEQUENCE 455 AA; 51566 MW; 12046F84A363F7A0 CRC64;
Query Match 41.7%; Score 758.5; DB 4; Length 455;
Best Local Similarity 44.0%; Pred. No. 8.7e-60;
Matches 150; Conservative 61; Mismatches 91; Indels 39; Gaps 5;
QY 1 MEKYEKLANTGEGSYGVFKCRNKTSGQVAVKKVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEWYETLGRVGEYSYGVWCKRKSQGRVAIKFLSEDDDDKWKIARIRKILKQLRH 59
QY 61 PNLVNLIEVFRKRRKHVLFVEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
DB 60 ENLVNLIEVFRQKKIHLVFEFIDHTVLDELQHYCHGLESKRLKYLQILRAIDYLSN 119
QY 121 NCIHRIKPEKNTLITKQGIKICDGFQAQILIPGDAYTDVYATWYRAPELLVGDTOYG 179
DB 120 NIHRIKPEKNTLVSGGVVKKLDCDFGFAITLAAPGDIYDVYATWYRAPELLVGDTSYG 179
QY 180 SSVDIWAIGCVFAELLTGQPLWPKGSDVDQVLYIIRTLVETGFRHVDQAGLELLTSSDP 239
DB 180 KPDVWALGCVLWMLGQPLPFGESDIQLHHITCL----- 217
QY 240 AVASQAGITGKLIPIHQSIKSNFFHGSIPEPEDMETLEEKFSQVHFVALNFKMCKL 299
DB 218 -----GNLSPLHLQNFPSKPIFAGVLPVQVHPKRNARKYPKLNGLLADIVHACL 267
QY 300 KKNPDRLTCSQLSESSYF--DSFOE---AQIKRKARNEGR 335
DB 268 QIDPADRISSDLLHHEYFTRDGFIEKFMPELKAQLQEAQ 308
RESULT 10
Q81VM4 PRELIMINARY: PRT: 592 AA.
AC Q81VM4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to cyclin-dependent kinase-like 3.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
```

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC041799; AHA1799.1; --
KW Kinase; Cyclin.
SQ SEQUENCE 592 AA, 67513 MW, 28:AFS8906E87697 CRC64;

Query Match 41.7%; Score 758.5; DB 4; Length 592;

Best Local Similarity 44.0%; Pred. No. 1.3e-59;

Matches 150; Conservative 61; Mismatches 91; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

DB 1 MEMYETLKGVEGSGYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 59

QY 61 PNLVNLIEVFRKRKMHVLFYCYCHTLLNELRNPNGVADGV:KSV:WQTLQALNFCFHH 123

DB 60 ENLVNLIEVFRKRKMHVLFYCYCHTLLNELRNPNGVADGV:KSV:WQTLQALNFCFHH 119

QY 121 NCIRHDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDVATRWYRAPELVGDTQYGS 179

DB 120 NIHRDIKPNILVSSQSGITK:ICDFGFAQIL-IPGDAYTDVATRWYRAPELVGDTQYGS 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPKGSVDQLYLIRTLVETGFRHVDQAGLELLTSSDPP 239

DB 180 KPDVIALGCMVIEATGNPVLPSDDLDLHLKLVK----- 217

QY 240 AVASAGITKGLIPRHOSIFKSNQFFHGISTEPEDEMETLEEKFSVDVHPVALNFMKGL 299

DB 218 -----GNLSPLQNIFSKSPFAGVLPVQHPKARKKYPKXGLGLADIVHACJ 267

QY 300 KNPDDRLTCSQLESSEYF--DSFQ-----AQIKSKARNEGR 335

DB 268 QIDPADRISSDLJHHEVFTREDGFIEKMPKELKAKJLCEAK 308

RESULT 11

QYBMG2 PRELIMINARY; PRT: 657 AA.

AC QYBMG2; "Trypanosoma brucei putative MAP kinase";

DT 01-JUN-2001 (TREMBLrel. 17, Created);

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update);

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update);

DE Putative MAP kinase.

OS Trypanosoma brucei.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

OX NCBI_TaxID=5691;

RN [1]

RP SEQUENCE FROM N.A.

RA Sarkar M., Matthews K.;

RT "Trypanosoma brucei putative MAP kinase";

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

CC -1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AF326965; AAG49599.1; --

DR HSSP: P24941; I838.

DR InterPro: IPRO01392; HLH basic.

DR InterPro: IPRO00719; Prot_kinase.

DR Pfam: PF00069; pkinase.

DR ProDom: PD000601; Prot_kinase; 1.

DR SMART: SM00220; S_TK; 1.

DR PROSITE: PS00038; HLH 1; 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 657 AA; 72116 MW; C93FA3C2D80F0619 CRC64;

Query Match 41.5%; Score 755; DB 5; Length 657;

Best Local Similarity 43.4%; Pred. No. 3e-59;

Matches 138; Conservative 67; Mismatches 81; Indels 32; Gaps 1;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

DB 1 MDAYETLGLGEGTYGVVFKARHRTSRIVAKYKQAEEDDDHVRKTSLEVRVRLKQLH 60


```
Db 180 KPVDMALGCKYLIEMATGHPFLPSSDDLHLKIVLVK----- 217
QY 240 AVASQAGITGKLIPIRHOS:FKSNGFFHGHSIPEPEDEMETLEEKFSQVHPVALNFMKGL 299
Db 218 -----GNLTPLHLNIFSKSPIFAGVLPQVQHPKTAFFKPKYKPLNGLLADIVHACL 267
QY 300 KKNPDRJCTCSOLLESSYF--DSFOEQ-----IKKARNEGNNRRSQ 34:
Db 268 QIDPAER:SSDILLRHDYPTROGFIEKKWTETKGPSSKSSSLRPGKAKEMSQTRRSQ 325

RESULT 13
Q3GR79 PRELIMINARY; PRT; 1106 AA.
AC Q3GR79;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 23, Last annotation update)
DE Putative mitogen-activated protein kinase 6.
GN MPK6
OS Leishmania mexicana.
OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishman.a.
OX NCBI_TaxID=5665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MNYC/BZ/62/M379;
RA Wiese M., Wang C., Goercke I.;
RT "Identification of mitogen-activated protein (MAP) kinases from Leishmania mexicana."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ293284; CAC07960.1;
DR HSSP; P24941; B38.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1106 AA; 118884 MW; 24AC26F5C209923E CRC64;

Query Match 40.9%; Score 743.5; DB 5; Length 1106;
Best Local Similarity 43.7%; Pred. No. 6.8e-58;
Matches 145; Conservative 62; Mismatches 88; Indels 37; Gaps 3;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVVKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEAYETLGILGEGTYGVVVKASRVTKLVAIKRFKQTEQDEHVEKTSREVRMLQLLQH 60
QY 61 ENLVNLIIEFRKRMHLVFEYCDHTLNELEERNNGVADGVKSVLWQTLQALNFCIH 120
Db 61 PVVILEDFVREGKLYLVFEIDHTILQLLESITRGRHRELRYTYQLLEG:EFCHNQ 120
QY 121 NCIHRIKPENILITKGGIK:CDGFGAQLIPGDAYTVATRAYRAPELLVGDTOYGS 180
Db 121 NIHRDVPENVL:DESGLLKLCDFGRQSAKGYIDYATRAYRAPELLVGDVAYGK 180
QY 181 SVDIWAIGCVPAELITGQPLWFGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSOPFA 240
Db 181 PVDVWALGCMFAELSDGQPLFGESDLDGLCLIMOT----- 216
QY 241 VASQAGITGKLIPIRHOS:FKSNGFFHGHSIPEPEDEMETLEEKFSQVHPVALNFMKGLK 300
Db 217 -----CGPVFQRLVFIEXHNPUNYNGISFPHTDILYTLKDRYHRESNDWIEFLSSCLH 268
QY 301 KNPDRJCTCSOLLESSYF--DSFQ---EAQIK 327
Db 269 TQPAQLTCTELMELPYFTRDGRFRYEALR 300
```

```
RESULT 14
Q3BKR2 PRELIMINARY; PRT; 457 AA.
AC Q3BKR2;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Serine/threonine kinase NKIATRE alpha homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RC EXBL; AK050990; BAC34498.1;
SQ SEQUENCE 457 AA; 51900 MW; 6PCDD36F5152B61 CRC64;

Query Match 40.7%; Score 741.5; DB 11; Length 457;
Best Local Similarity 42.8%; Pred. No. 3e-58;
Matches 146; Conservative 64; Mismatches 92; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVVKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEAYETLGKVGESYGVVVKCKKDKTGRIVAIIKPYE-KPEKSVNKIATREIKFLKQFRH 59
QY 61 ENLVNLIIEFRKRMHLVFEYCDHTLNELEERNNGVADGVKSVLWQTLQALNFCIH 120
Db 60 ENLVNLIIEFRKRMHLVFEYCDHTLNELEERNNGVADGVKSVLWQTLQALNFCIH 119
QY 121 NCIHRIKPENILITKGGIK:CDGFGAQLIPGDAYTVATRAYRAPELLVGDTCYV 179
Db 120 NIHRDVPENILVSGSITKLCDFGARTLAAPGDVYTVATRAYRAPELLVGDTCYV 179
QY 180 SVDIWAIGCVPAELITGQPLWFGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSOPP 239
Db 180 KPVDMALGCKYLIEMATGHPFLPSSDDLHLKIVLVK----- 217
QY 240 AVASQAGITGKLIPIRHOS:FKSNGFFHGHSIPEPEDEMETLEEKFSQVHPVALNFMKGL 299
Db 218 -----GNLTPLHLNIFSKSPIFAGVLPQVQHPKTAFFKPKYKPLNGLLADIVHACL 267
QY 300 KKNPDRJCTCSOLLESSYF--DSFOEQ---AQIKRKARNEG 335
Db 268 QIDPAERTSDDLRLRHDYPTROGFIEKFIPELRAKLQLEAK 308

RESULT 15
Q3BL49 PRELIMINARY; PRT; 585 AA.
AC Q3BL49;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Serine/threonine kinase NKIATRE alpha homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
RL
```

```
RL Nature 420:563-573(2002?);
DR EMBL: AK046394; BAC32701.1; -; A1867A2A6CDF5A2 CRC64;
SQ SEQUENCE 585 AA; 66530 MW;  A1867A2A6CDF5A2 CRC64;

Query Match      40.6%; Score 738.5; DB 11; Length 585;
Best Local Similarity 42.5%; Pred. No. 7.9e-58;
Matches 145; Conservative 65; Mismatches 92; Indels 39; Gaps 5;

QY 1 YEKYEKAKTCEGSGYGVYFKCRNKTSQGVAVKVKFVSEDDPVYKKIALREIRMLKQLKH 60
DB 1 MEMYETLQKVGQSGYGVYFKCRNKTSQGVAVKVKFVSEDDPVYKKIALREIRMLKQLKH 59

QY 61 PNLVNLIEVFRKRRKQKHLVFEYCYDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
DB 60 ENLVNLIEVFRKRRKQKHLVFEYCYDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 119

QY 121 NCIIHRDIKPENLITKQGIKICDFGFAQIL-IPGDAYTDYVATRYRAPELLVGTQYG 179
DB 120 NIIHRDIKPENLIVSQSGITKLCDFGFARTLAAPGDVYTDYVATRYRAPELVKDTSYG 179

QY 180 SSVETIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVE*GFRHVDQAGLELLTSSDPP 239
DB 180 KPVDIWALGCMIEWATGHPFLPSSOLDLLHKVLV----- 217

QY 240 AVASOSAGITKLIPIRHOS*FKSNGFFHGIS*PEPEDMETLEKFSVHPVVALNFMKGCL 299
DB 218 -----GNLTPHLHNI*FKSPIFAGVVLPOVQH*KTARKKYPK*NGLLADIVHACL 267

QY 300 KMN*PDRLLTCSLLESSYF--DSFOE---AQIKSKARNEGR 335
DB 268 QIDPAERTSSTCLLRHDYFTRDGFIEKFIPELRAKLLQEA 308
```

Search completed: November 13, 2003, 14:20:36
Job time : 35.0227 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CorpuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:12:21 ; Search time 36.1631 Seconds
(without alignments)
1382.592 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1662

Sequence: 1 MEKYEKLAKTGGSGYGVFK.....RKARNEGRRRRQQVJPLKS 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2030000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1662	100.0	315	AAE00495	Human kinase #6
2	1639	98.6	342	AAE19152	Human kinase polyp
3	1638	98.6	324	AAE00492	Human kinase #3.
4	1636	98.4	347	AAE00494	Human kinase #5.
5	1619	97.4	360	AAJ03525	Human protein: kina
6	1612	97.0	356	AAE00491	Human kinase #2.
7	1296.5	78.0	296	AAE65643	Novel protein kina
8	1188.5	71.5	358	AAE62954	Human polypeptide
9	1149	69.1	247	AAE65642	Novel protein: kina

10	1140	68.6	228	23	AAG78547	Human kinase 14257
11	1027.5	61.8	392	22	ABB63118	Drosophila melanog
12	967.5	58.2	566	21	ABU90724	Rabbit KKIAMRE kin
13	948	57.0	197	24	ABU11689	Human MDT1 polypep
14	937.5	56.4	362	24	ABP96087	Human protein kina
15	791	47.6	187	22	AAE00490	Human kinase #1.
16	791	47.6	198	22	AAE00493	Human kinase #4.
17	784.5	47.2	455	21	AAU70126	Human lost in leuk
18	784.5	47.2	591	22	AAU03524	Human protein kina
19	763.5	45.9	457	21	AAU70125	Rat lost in leukae
20	763.5	45.9	505	21	AAU70124	Rat lost in leukae
21	722.5	42.9	154	24	ABU05290	Human diagnostics
22	656	39.5	534	22	AAE65645	Novel protein kina
23	582	35.0	333	22	AAE39276	Human polypeptide
24	574	34.5	352	22	AAE41062	Human polypeptide
25	570	34.3	298	20	AAW95689	Homo sapiens Cdc2
26	570	34.3	544	19	AAW85028	CDK2-green fluorese
27	570	34.3	544	19	AAW85029	Green fluorescent
28	569	34.2	224	22	ABG06142	Novel human diagno
29	568.5	34.2	294	20	AAW95690	Oryza sativa Cdc2
30	568.5	34.2	297	23	ABE79004	Human 10ckshsl pro
31	566.5	34.1	294	21	AAG33356	Zea mays protein f
32	565.5	34.0	294	21	AAG40809	Zea mays protein f
33	561.5	33.8	294	21	AAG11214	Arabidopsis thalia
34	561.5	33.8	294	21	AAG54042	Arabidopsis thalia
35	561.5	33.8	350	21	AAE11213	Arabidopsis thalia
36	557	33.5	297	20	AAW95687	Rattus norvegicus
37	556	33.5	297	20	AAW95688	Homo sapiens Cdc2
38	556	33.5	297	22	AAE67434	Amino acid sequenc
39	556	33.5	297	22	AAE67613	Amino acid sequenc
40	556	33.5	297	24	ABR47405	Breast cancer asso
41	556	33.5	297	24	AAE34491	Human CDC2 protein
42	556	33.5	297	24	ABU56522	Lung cancer-associ
43	555.5	33.4	300	20	AAW95679	Pneumocystis carin
44	555	33.4	270	20	AAU43932	Human protein kina
45	554	33.3	298	21	AAU52184	Cyclin-dependent k

ALIGNMENTS

RESULT 1

AAE00495
ID AAE00495 standard; Protein; 315 AA.

AC AAE00495;

DT 19-JUN-2001 (first entry)

DE Human kinase #6.

DE Human; kinase; Gene therapy; bioreactor; mental disorder;

KW biological disorder.

XX Homo sapiens.

XX WO200123579-A1.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26621.

XX 28-SEP-1999; 99US-0156511.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX WPI; 2001-266166/27.

XX N-PSDB; AAD03817.

PT New isolated human kinase polynucleotide useful for generating

PT antibodies, as reagents in diagnostic assays and for screening for

compounds useful for treating mental, biological or medical diseases -

Claim 3; Page 33-34; 38pp; English.

The present sequence is novel human protein (NHP) known as human kinase. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.

Sequence 315 AA;

Query Match 100.0%; Score 1662; DB 22; Length 315;
 Best Local Similarity 100.0%; Pred. No. 1.9e-168;
 Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGQVAVKXFESEDDPVVKKIALREIMLKQLKH 60
 DB 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGQVAVKXFESEDDPVVKKIALREIMLKQLKH 60
 QY 61 PNLNLIIEVFRKRGHILVFEYCDHTLLNELRNPNGVAGGVKSVLWQTLQALNFCIH 120
 DB 61 PNLNLIIEVFRKRGHILVFEYCDHTLLNELRNPNGVAGGVKSVLWQTLQALNFCIH 120
 QY 121 NCIHRIKPEMLITKQGIKICDFGFAQIIPGCAVTDVATRWYRAPELLVSDTQYGS 180
 DB 121 NCIHRIKPEMLITKQGIKICDFGFAQIIPGCAVTDVATRWYRAPELLVSDTQYGS 180
 QY 181 SVDWAICGVFAELLTGQPLWFGKSDVDQLYLTARTGKLIPIHQKGFPHGISIP 240
 DB 181 SVDWAICGVFAELLTGQPLWFGKSDVDQLYLTARTGKLIPIHQKGFPHGISIP 240
 QY 241 EPEDMETLEEFSDVHPVALNFMKGLKMMNPDRLTCSQLESYFDSFQEAQIKKARN 300
 DB 241 EPEDMETLEEFSDVHPVALNFMKGLKMMNPDRLTCSQLESYFDSFQEAQIKKARN 300
 QY 301 EGRNRRRQCVLPKLS 315
 DB 301 EGRNRRRQCVLPKLS 315

RESULT 2
 AAEL9152
 IF AAEL9152 standard; Protein; 342 AA.
 XX AAEL9152;
 XX
 XX
 XX 21-MAY-2002 (first entry)
 XX Human Kinase polypeptide (PKIN-10).
 XX
 KW Human; kinase polypeptide; PKIN-10; gene therapy; Addison's disease;
 KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
 KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
 KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;

cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer; cholestasis; cardiac; cardiovascular disorder; Niemann-Pick's disease; lipid disorder; fatty liver; Gaucher's disease; myocardial infarction; drug screening; transgenic animal; anti-inflammatory; hepatotropic; hypotensive; anti-HIV; enzyme.

XX Homo sapiens.

PH Key Location/Qualifiers
 FH Domain 4..286
 FT /note= "Eukaryotic protein kinase domain"

XX MC200228399-A2.
 XX 31-JAN-2002.
 XX 20-JUL-2001; 2001WO-US23092.
 XX 21-JUL-2000; 2000US-220038P.
 XX 28-JUL-2000; 2000US-222112P.
 XX 04-AUG-2000; 2000US-222831P.
 XX 11-AUG-2000; 2000US-224729P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX (THOR/) THORNTON M.
 XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Wallia NK, Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR, Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lai P, Ding L, Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DM, Greenwald SR, Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;
 XX WPI; 2002-206083/26.
 XX N-PSDB; RAD30557.
 DR New human kinase polypeptide, useful in diagnosis, prevention and treatment of cancer, immune disorder, growth and developmental disorder, cardiovascular disorder and lipid disorder -
 XX Claim 1; Page 147-148; 196pp; English.
 XX The present invention relates to an isolated human kinase polypeptide (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is useful for diagnosing, treating and preventing cancer (e.g., leukaemia, lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's disease, rheumatoid arthritis), a growth and developmental disorder (e.g., bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial infarction), and a lipid disorder (e.g., fatty liver, cholestasis, Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of drug screening techniques and to analyse the proteome of a tissue or cell type. PKIN is useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals and as hybridisation probes for mapping naturally occurring genomic sequences. PKIN is useful in southern or northern analysis, dot blot or other membrane-based technologies, in PCR technologies, in dipstick, pin, multiformat enzyme linked immunosorbent (ELISA)-like assays and in microarrays utilising fluids or tissues from patients to detect altered PKIN expression. The present sequence is human PKIN-10.
 XX Sequence 342 AA;
 XX

Query Match 98.6%; Score 1639; DB 23; Length 342;
 Best Local Similarity 99.0%; Pred. No. 6e-166;
 Matches 312; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGQVAVKXFESEDDPVVKKIALREIMLKQLKH 60
 |||
 |||

Db 1 MEKYELANTGEGSYGVVFKCRNKTSQQVAVVKKFVSEDDPVVKKIALREIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNPGVADGVKSVLACTLOALNFCHH 120
 Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNPGVADGVKSVLACTLOALNFCHH 120
 QY 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRYRAPELLVGTQYGS 180
 Db 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRYRAPELLVGTQYGS 180
 QY 181 SVDIWAIGCVFAELITGQPLWPKSDVDOLYLIIRTLGKLI PRHQSIFKSNQFFHGISIP 240
 Db 181 SVDIWAIGCVFAELITGQPLWPKSDVDOLYLIIRTLGKLI PRHQSIFKSNQFFHGISIP 240
 QY 241 EPEDMETLEEKFSVHPVAFKMGCLKNPDRLTCSQLLESSYFDSFQEAQIKRKARN 300
 Db 241 EPEDMETLEEKFSVHPVAFKMGCLKNPDRLTCSQLLESSYFDSFQEAQIKRKARN 300
 QY 301 EGNRRRQV--QVLPL 313
 Db 301 EGNRRRQV--QVLPL 313

RESULT 3

AAECG492
 ID AAE00492 standard; Protein; 324 AA.

XX AC AAE00492;

XX DT 19-JUN-2001 (first entry);

XX DE Human kinase #3.

XX KW Human; kinase; gene therapy; bioreactor; mental disorder;

XX KW biological disorder.

XX OS Homo sapiens.

XX PN WO200123579-A1.

XX PD 05-APR-2001.

XX PF 27-SEP-2000; 2000WO-US26621.

XX PR 28-SEP-1999; 99US-0156511.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX DR WPI; 2001-266166/27.

XX DR N-PSDB; AAD03814.

XX PT New isolated human kinase polynucleotide useful for generating

XX PT antibodies, as reagents in diagnostic assays and for screening for

XX PT compounds useful for treating mental, biological or medical diseases

XX PS Disclosure; Page 30; 38pp; English.

XX CC The present sequence is novel human protein

XX CC (NHP) known as human kinase. The human kinases share structural

XX CC similarity with animal kinases, more particularly serine or

XX CC threonine protein kinases. Human kinase cDNA is useful for the

XX CC detection of mutant human kinase for the diagnosis of disease,

XX CC and also as a therapeutic. It is useful for screening drugs

XX CC effective in the treatment of symptomatic or phenotypic

XX CC manifestations perturbing the normal function of NHP in the

XX CC body. The NHP nucleotide sequences are useful for generation of

XX CC antibodies, as reagents in diagnostic assays, for the

XX CC identification of other cellular gene products related to human

XX CC kinases, and as reagents in assays for screening compounds that

XX CC are useful for treating mental, biological or medical disorders.

XX CC NHP oligonucleotides are used as probes. The labeled NHP probes

CC are useful for screening human genomic library for identifying
 CC polymorphisms and as primers in amplification assays to detect
 CC mutations within the exons, introns and splice sites that can
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct
 CC encoding NHP products are used to genetically engineer cells
 CC in vivo that functions as bioreactors in the body delivering a
 CC continuous supply of NHP to the body. Nucleotide constructs
 CC encoding functional NHPs are used in gene therapy for the
 CC modulation of NHP expression.

XX SC Sequence 324 AA;

Query Match 98.6%; Score 1639; DB 22; Length 324;

Best Local Similarity 100.0%; Pred. No. 7.1e-166;

Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYELANTGEGSYGVVFKCRNKTSQQVAVVKKFVSEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEKYELANTGEGSYGVVFKCRNKTSQQVAVVKKFVSEDDPVVKKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNPGVADGVKSVLACTLOALNFCHH 120

Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNPGVADGVKSVLACTLOALNFCHH 120

QY 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRYRAPELLVGTQYGS 180

Db 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRYRAPELLVGTQYGS 180

QY 181 SVDIWAIGCVFAELITGQPLWPKSDVDOLYLIIRTLGKLI PRHQSIFKSNQFFHGISIP 240

Db 181 SVDIWAIGCVFAELITGQPLWPKSDVDOLYLIIRTLGKLI PRHQSIFKSNQFFHGISIP 240

QY 241 EPEDMETLEEKFSVHPVAFKMGCLKNPDRLTCSQLLESSYFDSFQEAQIKRKARN 300

Db 241 EPEDMETLEEKFSVHPVAFKMGCLKNPDRLTCSQLLESSYFDSFQEAQIKRKARN 300

QY 301 EGNRRRQV 310

Db 301 EGNRRRQV 310

RESULT 4

AAECG492

ID AAE00494 standard; Protein; 347 AA.

XX AC AAE00494;

XX DT 19-JUN-2001 (first entry)

XX DE Human kinase #5.

XX KW Human; kinase; gene therapy; bioreactor; mental disorder;

XX KW biological disorder.

XX OS Homo sapiens.

XX PN WO200123579-A1.

XX PD 05-APR-2001.

XX PF 27-SEP-2000; 2000WO-US26621.

XX PR 28-SEP-1999; 99US-0156511.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX DR WPI; 2001-266166/27.

XX DR N-PSDB; AAD03816.

XX PT New isolated human kinase polynucleotide useful for generating

XX PT antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases -
 XX Claim 2; Page 32-33; 38pp; English.

XX The present sequence is novel human protein
 CC (NHP) known as human kinase. The human kinases share structural
 CC similarity with animal kinases. More particularly serine or
 CC threonine protein kinases. Human kinase cDNA is useful for the
 CC detection of mutant human kinase for the diagnosis of disease,
 CC and also as a therapeutic. It is useful for screening drugs
 CC effective in the treatment of symptomatic or preneoplastic
 CC manifestations perturbing the normal function of NHP in the
 CC body. The NHP nucleotide sequences are useful for generation of
 CC antibodies, as reagents in diagnostic assays, for the
 CC identification of other cellular gene products related to human
 CC kinases, and as reagents in assays for screening compounds that
 CC are useful for treating mental, biological or medical disorders.
 CC NHP oligonucleotides are used as probes. The labeled NHP probes
 CC are useful for screening human genomic library for identifying
 CC polymorphisms and as primers in amplification assays to detect
 CC mutations within the exons, introns and splice sites that can
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct
 CC encoding NHP products are used to genetically engineer cells
 CC in vivo that functions as bioreactors in the body delivering a
 CC continuous supply of NHP to the body. Nucleotide constructs
 CC encoding functional NHPs are used in gene therapy for the
 CC modulation of NHP expression.

XX Sequence 347 AA;

Query Match 98.4%; Score 1636; DB 22; Length 347;
 Best Local Similarity 90.8%; Pred. No. 1.3e-165;
 Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
 QY 1 MEKYEKLTGEGSGVGVFKCNKTSQGVAVKVFSEDDPVVKKIALRIRMLKQLKH 60
 DB 1 MEKYEKLTGEGSGVGVFKCNKTSQGVAVKVFSEDDPVVKKIALRIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKRKHVLFVEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
 DB 61 PNLVNLIEVFRKRKHVLFVEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
 QY 121 NCIHRRDIKPENILITKQGIKICDFGFAQILPGDAYTYVATRYRAPELLVCTOYGS 180
 DB 121 NCIHRRDIKPENILITKQGIKICDFGFAQILPGDAYTYVATRYRAPELLVCTOYGS 180
 QY 181 SVDIWAIGCVFAELLTGQPLWPKSKSDVDQGLYILRTLVETGPRHVDQAGLELLTSSOFFA 240
 DB 181 SVDIWAIGCVFAELLTGQPLWPKSKSDVDQGLYILRTLVETGPRHVDQAGLELLTSSOFFA 240
 QY 218 -----GKLIPIHCSIFKSGNFFHGHSIPEPEDEMETLEEKSDVHPVALNFMKGLK 269
 DB 241 VASQAGITGKLIPIHQSIKSGNFFHGHSIPEPEDEMETLEEKSDVHPVALNFMKGLK 300
 QY 269 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRNRQQVLPKLS 315
 DB 301 XNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRNRQQVLPKLS 347

RESULT 5
 AAU03525

ID AAU03525 standard; Protein; 360 AA.

XX AAU03525;

AC AAU03525;

XX 12-SEP-2001 (first entry)

DE Human protein kinase #25.

XX Human; protein kinase: PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder.

XX Homo sapiens.
 CS WO200138503-A2.
 PN 31-MAY-2001.
 XX 22-NOV-2000; 2000WO-US32085.
 XX 24-NOV-1999; 99US-0167482.
 XX (SUGS-) SUGEN INC.
 PA Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 FI N-PSDB; AA506725.
 DR WPI; 2001-343950/36.
 XX N-PSDB; AA506725.
 XX Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 PS Claim 7; Figure 2; 43pp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.

XX Sequence 360 AA;

Query Match 97.4%; Score 1619; DB 22; Length 360;
 Best Local Similarity 93.4%; Pred. No. 8.9e-164;
 Matches 311; Conservative 2; Mismatches 0; Indels 20; Gaps 2;

QY 1 MEKYEKLTGEGSGVGVFKCNKTSQGVAVKVFSEDDPVVKKIALRIRMLKQLKH 60
 DB 1 MEKYEKLTGEGSGVGVFKCNKTSQGVAVKVFSEDDPVVKKIALRIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKRKHVLFVEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
 DB 61 PNLVNLIEVFRKRKHVLFVEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
 QY 121 NCIHRRDIKPENILITKQGIKICDFGFAQILPGDAYTYVATRYRAPELLVCTOYGS 162
 DB 121 NCIHRRDIKPENILITKQGIKICDFGFAQILPGDAYTYVATRYRAPELLVCTOYGS 180
 QY 163 TRWYRAPELLVCTOYGSVDIWAIGCVFAELLTGQPLWPKSKSDVDQGLYILRTLVETGPRH 222
 DB 161 TRWYRAPELLVCTOYGSVDIWAIGCVFAELLTGQPLWPKSKSDVDQGLYILRTLVETGPRH 240
 QY 223 RHQSIFKSGNFFHGHSIPEPEDEMETLEEKSDVHPVALNFMKGLKKNPDRLTCSQLLE 282
 DB 241 RHQSIFKSGNFFHGHSIPEPEDEMETLEEKSDVHPVALNFMKGLKKNPDRLTCSQLLE 300
 QY 283 SSYFDSFQEAQIKRKARNEGRRNRQQVLPKLS 313
 DB 301 SSYFDSFQEAQIKRKARNEGRRNRQQVLPKLS 333

```

RESULT 6
AAE00491
ID AAE00491 standard; Protein; 356 AA.
XX
XX AAE00491;
AC
NC
AC
DT 19-JUN-2001 (first entry)
DE
DE Human Kinase #2.
XX
XX Human; kinase; gene therapy; bioreactor; mental disorder;
XX biological disorder.
XX
XX Homo sapiens.
OS
OS
PN WO200123579-A1.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26621.
XX
XX 28-SEP-1999; 99US-0156511.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Doncho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
XX MPI; 2001-266166/27.
XX
XX N-PSDB; AAO03813.
XX
XX
XX New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for
XX compounds useful for treating mental, biological or medical diseases
XX
XX Disclosure; Page 28-29; 38pp; English.
XX
XX
XX The present sequence is novel human protein
XX (NHP) known as human kinase. The human kinases share structural
XX similarity with animal kinases, more particularly serine or
XX threonine protein kinases. Human kinase cDNA is useful for the
XX detection of mutant human kinase for the diagnosis of disease,
XX and also as a therapeutic. It is useful for screening drugs
XX effective in the treatment of symptomatic or phenotypic
XX manifestations perturbing the normal function of NHP in the
XX body. The NHP nucleotide sequences are useful for generation of
XX antibodies, as reagents in diagnostic assays, for the
XX identification of other cellular gene products related to human
XX kinases, and as reagents in assays for screening compounds that
XX are useful for treating mental, biological or medical disorders.
XX NHP oligonucleotides are used as probes. The labelled NHP probes
XX are useful for screening human genomic library for identifying
XX polymorphisms and as primers in amplification assays to detect
XX mutations within the exons, introns and splice sites that can
XX be used in diagnostics and pharmacogenomics. Nucleotide construct
XX encoding NHP products are used to genetically engineer cells
XX in vivo that functions as bioreactors in the body delivering a
XX continuous supply of NHP to the body. Nucleotide constructs
XX encoding functional NHPs are used in gene therapy for the
XX modulation of NHP expression.
XX
XX Sequence 356 AA;
XX
XX Query Match 97.0%; Score 1612; DB 22; Length 356;
XX Best Local Similarity 90.6%; Ref. No. 4.9e-163;
XX Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
XX
XX 1 MEKYEKLTGSGYGVFKCNKTSQGVVAVKVFESDDCPVVKKIALREIRN-KQJKH 60
XX
XX 1 MEKYEKLTGSGYGVFKCNKTSQGVVAVKVFESDDCPVVKKIALREIRN-KQJKH 60
XX
XX 61 PNLVNLIIEVFRKRRKMHLYFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
XX

```

CC Chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 CC
 XX
 SQ Sequence 296 AA;
 Query Match 78.0%; Score 1296.5; DB 22; Length 296;
 Best Local Similarity 90.0%; Pred. No. 1.7e-129;
 Matches 243; Conservative 19; Mismatches 5; Indels 3; Gaps 2;
 QY 46 KVALREIRMLKQKHPNLVNLIEVRRKRKXHLVFEYCDHTLLNELRNPNQVADGVKS 105
 DB 1 KALREIRMLK-LKHPNLVNLIEVRRKRKXHLVFEYCDHTLLNELRNPNQVADGVKS 59
 QY 106 VLMQTLQALNFCIHNCIHRDKPENILITKGIKICDFGFAQLIPGDAYTVVATRW 165
 DB 60 VLMQTLQALNFCIHNCIHRDKPENILITKGIKICDFGFAQLIPGDAYTVVATRW 119
 QY 166 YRAPELLVGDQYQSSVDIWAICGVFAELTQOPLWPKSDVDQYLIIRTLGKLIPRHQ 225
 DB 120 YRAPELLVGDQYQSSVDIWAICGVFAELTQOPLWPKSDVDQYLIIRTLGKLIPRHQ 179
 QY 226 SJFKNSGFFPHGISIPEDMETLEKFSVDHPVALNFKGCLKXNPDRLTCSQLLESSY 285
 DB 180 SJFKNSGFFPHGISIPEDMETLEKFSVDHPVALNFKGCLKXNPDRLTCSQLLESSY 239
 QY 286 FQFOFAQIKRKARNEGRNRQ--QVLP 313
 DB 240 FQFOFAQIKRKARNEGRNRQ--QVLP 269
 RESULT 8
 ABP62954
 ID ABP62954 standard; Protein; 358 AA.
 XX
 AC ABP62954;
 XX
 DT 14-OCT-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 391.
 XX
 KW Human; vulnary; dermatological; neuroprotective; nootropic; cancer;
 KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
 KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
 KW burn; central nervous system disorder; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; immune disorder;
 KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WQ200218424-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 31-AUG-2003; 2001WO-US27093.
 XX
 PR 01-SEP-2000; 2000US-0654935.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
 PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
 XX
 DR WPI: 2002-383321/62.
 DR N-PSDB; ABQ93433.
 XX
 PT New polynucleotide and polypeptides, useful for treatment and diagnosis
 PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
 PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
 PT sclerosis, diabetes and allergies -
 XX
 PS Claim 20; SEQ ID NO 391; 284pp + Sequence Listing; English.
 XX

CC The invention relates to an isolated polynucleotide (I) comprising one of
 CC 245 sequences (AB093288-AB093532). Treating a condition comprising
 CC administering to a mammalian subject a composition comprising the protein
 CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
 CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
 CC (I) is useful for gene therapy of diseases and (II) can be used for
 CC therapeutic treatment. Diseases that may be treated include wound healing
 CC and tissue repair, burns, central nervous system disorders (e.g.
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
 CC sclerosis, diabetes and allergies.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 358 AA;
 Query Match 71.5%; Score 1188.5; DB 23; Length 358;
 Best Local Similarity 70.5%; Pred. No. 7.8e-118;
 Matches 213; Conservative 42; Mismatches 46; Indels 1; Gaps 1;
 QY 1 MEKYEKLAXTGSGYGVVFKCRNKTSGGVAVKFESEDDPWWKKIALREIRMLKQLKH 60
 DB 2 MEKYEKIGKIGSGYGVVFKCRNKTSGGVAVKFESEDDPWWKKIALREIRMLKQLKH 61
 QY 61 PNLVNLIEVRRKRKXHLVFEYCDHTLLNELRNPNQVADGVKSVLWQTLQALNFCIH 120
 DB 62 PNLVNLIEVRRKRKXHLVFEYCDHTLLNELRNPNQVADGVKSVLWQTLQALNFCIH 121
 QY 121 NCIHRDKPENILITKGIKICDFGFAQLIPGDAYTVVATRWYRAPPELLVGDQYQ 179
 DB 122 NCIHRDKPENILITKGIKICDFGFAQLIPGDAYTVVATRWYRAPPELLVGDQYQ 181
 QY 180 SSVDIWAICGVFAELTQOPLWPKSDVDQYLIIRTLGKLIPRHQSPKSGFFPHGISI 239
 DB 182 PPVDVWAICGVFAELTQOPLWPKSDVDQYLIIRTLGKLIPRHQSPKSGFFPHGISI 241
 QY 240 PPEDMETLEKFSVDHPVALNFKGCLKXNPDRLTCSQLLESSYFQFOFAQIKRKAR 299
 DB 242 PPEDMETLEKFSVDHPVALNFKGCLKXNPDRLTCSQLLESSYFQFOFAQIKRKAR 301
 QY 300 NE 301
 DB 302 NK 303
 RESULT 9
 AAB65642
 ID AAB65642 standard; Protein; 247 AA.
 XX
 AC AAB65642;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase, SEQ ID NO: 169.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pevic disease; multiple sclerosis; psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WQ200073469-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US14842.
 XX
 PS 28-MAY-1999; 99US-0136503.
 XX

CC The invention may also be of use as therapeutic agents in cardiovascular
CC disorders such as hypertension and coronary artery disease, and some
CC endothelial cell disorders, including psoriasis. The current
CC sequence represents a human kinase 14257 amino acid sequence.

XX Sequence 228 AA;
SQ Query Match: 68.6%; Score 1140; DB 23; Length 228;
Best Local Similarity 98.2%; Pred. No. 6e-113;
Matches 214; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MEKYELAKTGGSGYGVVFKCRNKTSGQVAVKXFESEDDPVVVKIALREIRMLKOLKH 60
DB 1 MGKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKXFESEDDPVVVKIALREIRMLKOLKH 60
QY 61 PNLVNLIEVFRKRKMHVLFVFCYCDHTLLNELERNPNVAGGVKSVLWQTLQALNFCIH 120
DB 61 PNLVNLIEVFRKRKMHVLFVFCYCDHTLLNELERNPNVAGGVKSVLWQTLQALNFCIH 120
QY 121 NCIHRRDIPENILITKQGIKICDFGFAQLIPGDAYTDVATRWYRAPELLVGDTCYGS 180
DB 121 NCIHRRDIPENILITKQGIKICDFGFAQLIPGDAYTDVATRWYRAPELLVGDTCYGS 180
QY 181 SVDIWAICGVFAELLTGQPLWFGKSDVDQLYLIIRTLG 218
DB 181 SVDIWAICGVFAELLTGQPLWFGKSDVDQLYLIIRTLG 218

RESULT 1:
AB863118
ID AB863118 standard; Protein; 392 AA.
XX AC AB863118;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 16146.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US03231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX P: Venter JC, Adams M, Li PWD, Myers EW;
XX WP: 2001-656860/75.
XX DR N-PSDB; AB07221.
XX PR New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure; SEQ ID NO 16146; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB016176-AB030511), expressed DNA
CC sequences (AB01840-AB016175) and the encoded proteins
CC (AB057737-AB072072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 392 AA;
SQ Query Match: 61.8%; Score 1027.5; DB 22; Length 392;
Best Local Similarity 58.4%; Pred. No. 1.4e-100;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;
QY 1 MEKYELAKTGGSGYGVVFKCRNKTSGQVAVKXFESEDDPVVVKIALREIRMLKOLKH 60
DB 1 MDREYKLSRLGEGSYGVVFKCRDRETCALVAVKXFESEDDPAIRKIALREIRLLKNLKH 60
QY 61 PNLVNLIEVFRKRKMHVLFVFCYCDHTLLNELERNPNVAGGVKSVLWQTLQALNFCIH 120
DB 61 PNLVNLIEVFRKRKMHVLFVFCYCDHTLLNELERNPNVAGGVKSVLWQTLQALNFCIH 120
QY 121 NCIHRRDIPENILITKQGIKICDFGFAQLIPGDAYTDVATRWYRAPELLVGDTCYGS 180
DB 121 GCLHRDIPENILITACQGVKLCDFGFARMLSPGENYTDVATRWYRAPELLVGDTCYGT 180
QY 181 SVDIWAICGVFAELLTGQPLWFGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNPFHGSIIP 240
DB 181 PVDVWAIGCLFAELVRGEALWPGKSDVDQLYLIIRTKLIGDLPRIHQIFGQKEYFKGITLP 240
QY 241 EPEDMETLEERF---SDVHPVALNFMKGLKMNPDRLTCSOLLESSYVDSF--QEAQIK 295
DB 241 VPPTLEPLEDKPAKSOQNPLTIDFLKKCLDKDPTKRWSCCKLTKHSYFDYIAKQRELE 300
QY 296 RKARNEGNNRRRQQV 310
DB 301 HVNSLEAANLRQQQL 315

RESULT 12
AAY90724
ID AAY90724 standard; Protein; 566 AA.
XX AC AAY90724;
XX DT 15-AUG-2000 (first entry)
XX DE Rabbit KIAMRE kinase SEQ ID NO:4.
XX KW Rabbit; KIAMRE kinase; learning-induced kinase; learning; memory;
XX KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;
XX KW identification.
XX OS Oryctolagus cuniculus.
XX PN WC2000020567-A2.
XX PD 13-APR-2000.
XX PF 01-OCT-1999; 99WO-US23010.
XX PR 02-OCT-1999; 98US-0102906.
XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX PT Thompson RF, Gomi H, Sun W;
XX WP: 2000-329332/28.
XX DR N-PSDB; AAA29745.
XX PR Novel learning induced kinase polynucleotides and polypeptides, useful
XX PT for the analysis of learning and memory, and for gene therapy -
XX PS Claim 1; Fig 4; 64pp; English.
XX CC The present sequence represents a learning-induced kinase, designated
XX CC KIAMRE kinase, which is isolated from rabbit brain tissue. KIAMRE

kinase is a cdc2-related kinase. The KXIAMRE kinase polynucleotides can be used to express recombinant protein for analysis, characterisation or therapeutic use, as markers for tissues in which the protein is preferentially expressed, as molecular weight markers or Southern gels, as chromosome markers or tags, to compare endogenous DNA sequences in patients to identify potential genetic disorders, as probes to hybridise and discover novel related sequences, as a source of PCR primers, and as an antigen to induce anti-DNA antibodies. The polypeptides can be used in assay to discover biological activity, to raise antibodies, as tissue markers, and to isolate correlative receptors or ligands. The polynucleotides may also be used for gene therapy for the treatment of disorders which are mediated by KXIAMRE kinase.

Query Match 58.2%; Score 967.5; DB 21; Length 566;
 Best Local Similarity 54.6%; Pred. No. 6e-24;
 Matches 177; Conservative 67; Mismatches 65; indels 15; Gaps 5

```

QY 1 MEKYEKLAKTGBGSGVGVFKCNKTSQGVAVKXFVESEDPVVVXKIALREIRMLKQLAH 60
D6 1 MEKYNLGLVGGSGYGMVWKCRKNDGSRIVA'KKFLESDZDVVVKXIANWE'KLLKQLAH 60
QY 61 PNLVNLIEVFRBKRKWHVFEVWCHTLLNELERNPNGVADGVIKSVLQTL'CALNE'CHIH 120
D6 61 ENLVNLLEVCKKKKRWYLVFVPHDITLDDLEF'FNGLDQV'QVKYLFQ'INGIF'CHSH 120
QY 121 NCIHROIKPENLITKQGI'IKICDFGPAQIL-IPGDAYTDVAVTRAVRAPELLAGDVOYG 179
D6 121 NIUHRDIKPEN'LVSQSGVWLKDFGFPARTLAAPGEVYTDVAVTRAVRAPELLVGDVKYG 180
QY 180 SSVDIWAIGCVFAELT'COPLWP'KQSDVDLYLI'RTLKGLIBRH'SI'FKSGN'GFPHGIS' 239
D6 181 KAVDVAIGCLVTEMLWGEPL'FGCSDIDOLY'LYRCLGNLIBRH'QEL'FVKPVP'AGVPL 240
QY 240 PPEPMETLEKSFSDVHPVA'NFVKKGCJKNPDRLTCS'OLLESSVYF--DSF-----QEA 292
D6 241 PEIKSEPLERRYPKLSBW'CLAKKCLHVJPD'KRPFCEALLHCH'DFFQVGDGAERFSQEL 300
QY 293 QIK--RKARN-----EGNRNRRQC 309
D6 301 QMKVQKADARNISLSKKSQNRKKE 324

```

RESULT 13
ABU11689
ID ABU11689 standard; Protein: 197 AA.

XX AC ABU11689;

DT 13-FEB-2003 (first entry)

Human MDDT polypeptide SEQ ID 636.

MDDT; human; disease detection and treatment molecule polypeptide;
 anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
 haemostatic; nephrotropic; antianaemic; antiproliferic; hepatotropic;
 gene therapy; protein replacement therapy; cell proliferative disorder;
 cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
 anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
 psoriasis; hepatitis.

XX
OS
Homo sapiens.

XX
PN
WO200279449-A2.XX
PD
10-OCT-2002.XX
PF 27-MAR-2002: 2002WO-US09944.XX
PR 28-MAR-2001 2001:5-279619D

PR 29-MAR-2001; 2001US-280067P.

```

PR 23-MAR-2001; 2001US-280068P.
PR 16-MAY-2001; 2001US-291280P.
PR 17-MAY-2001; 2001US-291829P.
PR 17-MAY-2001; 2001US-291843P.
PR 19-JUN-2001; 2001US-299428P.
PR 20-JUN-2001; 2001US-299776P.
PR 20-JUN-2001; 2001US-300001P.
XX
XX
(ENCY-) INCYTE GENOMICS INC.
XX
XX
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gierzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tucson O, Yap PE, Amshay SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
XX WFI: 2003-058431/05.
DR DR N-PSDB; ABX34679.
XX
XX
XX New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers
PT (c.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
F: c/ hepatitis -
XX
XX Claim 27; SEQ ID NO 636; 339pp + Sequence Listing: English.
XX

```

This invention describes a novel disease detection and treatment molecule polypeptide (MDPT) which has anti-inflammatory, immunosuppressive, osteopetastic, cytostatic, anti-HIV, haemostatic, nephrotropic, antianemic, antipsoriatic and hepatotropic activity. The polynucleotides and the polypeptides of the invention can be used for gene therapy, protein replacement therapy and are useful for treating a variety of diseases or conditions. These polypeptides or polynucleotides are particularly useful for diagnosing, treating or preventing cell proliferative disorders (e.g. cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's syndrome, inflammation, osteoporosis, thrombocytopaenia, psoriasis or hepatitis. ABU1450-ABU1845 represent the MDPT polynucleotides encoded by ABU1450-ABU1845, described in the disclosure of the invention. NCRT: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format from WIPO at [ftp.wipo.int/pub/published/pct_sequences](http://wipo.int/pub/published/pct_sequences).

AA	
Sequence	197 AA;

Query Match	57.0%;	Score 948;	DB 24;	Length 197;
Best Local Similarity	98.9%;	Pred. No. 1.5e-92;		
Matches 180;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps 0;				

[illegible]

QY 61 PNLVNLI EVFRKRKMH LVFEYCDHTLLNELERNP NGVADGVIKSVLWOTLOALNFCHZH 120

db 66 PNLVNLIIEVFRRKQMHVLFHEYCDHTLLGLERNPNGVADGVIKSVLWOTLOALNFECHIH 125

QY 121 NC1HRDIKPENILITKOGIIKICDFGFAOILIPGDAYTDYVATRWYRAPPELLVGDPTOYGS 180

Db 126 NC IHRDIK PENILITKOGIIKICDFGFAQILIPGDAYTDYAATRWYRAPELLVGDTQYGS 185

Qy 181 SV 182

QC
186 SV 187

RESULT 14

ABP96087
ID ABP96087 standard; Protein: 362 AA;

XX
AC ABP96087;

AC
XX
ABF 36067;

[illegible]

```

QY 1 MEKYEKLAKTGEYSYGVWFKCRNKTSQVAVVKKFVESEDDPVKKIALREIRMLKQLKH 60
DE 1 MEKYEKLAKTGEYSYGVWFKCRNKTSQVAVVKKFVESEDDPVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLLNELEARNPNGVADGVTKSVJWQTLQALNFCIH 120
DB 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLLNELEARNPNGVADGVTKSVJWQTLQALNFCIH 120
QY 121 NCIHRIKIPENILITKQGIKIICDFGFAQIL 151
DB 121 NCIHRIKIPENILITKQGIKIICDFGFAQIL 151

```

Search completed: November 13, 2003, 14:18:48
Job time : 37.1631 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:17:22 ; Search time 14.2749 Seconds
(without alignments)
933.660 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1662

Sequence: 1 MEKYEKAKTGEGSYGVVFK.....RKARNEGRRRQQVLP:KS 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4231085a residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata1/1/aa/5A_COMB.psp:*

2: /cgn2_6/ptodata1/1/aa/5B_COMB.psp:*

3: /cgn2_6/ptodata1/1/aa/6A_COMB.psp:*

4: /cgn2_6/ptodata1/1/aa/6B_COMB.psp:*

5: /cgn2_6/ptodata1/1/aa/PTUS_COMB.psp:*

6: /cgn2_6/ptodata1/1/aa/backfiles1.psp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1183.5	71.2	358	4	US-09-411-628-11	Sequence 11, Appl
2	967.5	58.2	566	4	US-09-411-628-4	Sequence 4, Appl
3	941.5	56.6	493	4	US-09-411-628-10	Sequence 10, Appl
4	570	34.3	298	2	US-08-874-347-25	Sequence 25, Appl
5	570	34.3	298	2	US-08-969-106-2	Sequence 2, Appl
6	570	34.3	298	3	US-09-093-522-25	Sequence 25, Appl
7	570	34.3	298	4	US-09-338-125-2	Sequence 2, Appl
8	570	34.3	298	4	US-09-266-225D-14	Sequence 14, Appl
9	570	34.3	544	4	US-09-417-197-113	Sequence 113, App
10	570	34.3	544	4	US-09-417-197-115	Sequence 115, App
11	569	34.2	298	4	US-09-411-628-13	Sequence 13, Appl
12	568.5	34.2	294	2	US-08-874-347-26	Sequence 26, Appl
13	568.5	34.2	294	3	US-09-093-522-26	Sequence 26, Appl
14	567	34.1	298	4	US-09-457-040B-29	Sequence 29, Appl
15	565	34.0	297	1	US-08-176-620A-16	Sequence 16, Appl
16	565	34.0	297	2	US-08-461-985-16	Sequence 16, Appl
17	557	33.5	297	2	US-08-874-347-23	Sequence 23, Appl
18	557	33.5	297	3	US-09-093-522-23	Sequence 23, Appl
19	556	33.5	297	2	US-08-874-347-24	Sequence 24, Appl
20	556	33.5	297	3	US-09-093-522-24	Sequence 24, Appl
21	556	33.5	297	4	US-09-411-628-12	Sequence 12, Appl
22	555.5	33.4	300	2	US-08-874-347-10	Sequence 10, Appl
23	555.5	33.4	300	3	US-09-093-522-10	Sequence 10, Appl
24	555	33.4	270	2	US-07-857-2248-31	Sequence 31, Appl
25	551	33.2	274	1	US-08-318-947A-20	Sequence 20, Appl
26	551	33.2	274	2	US-08-795-303-20	Sequence 20, Appl
27	544	32.7	292	1	US-08-154-915-2	Sequence 2, Appl

28	544	32.7	292	2	US-08-464-517-38	Sequence 38, Appl
29	544	32.7	292	2	US-08-246-361A-38	Sequence 38, Appl
30	544	32.7	292	3	US-08-463-772-38	Sequence 38, Appl
31	544	32.7	292	5	PCT-US93-09945-2	Sequence 2, Appl
32	537	32.3	297	3	US-08-932-787B-21	Sequence 21, Appl
33	537	32.3	297	3	US-08-932-012C-21	Sequence 21, Appl
34	537	32.3	297	4	US-08-888-818C-21	Sequence 21, Appl
35	518.5	31.2	297	2	US-08-874-347-22	Sequence 22, Appl
36	518.5	31.2	297	3	US-09-093-522-22	Sequence 22, Appl
37	511	30.7	274	1	US-07-857-224B-30	Sequence 30, Appl
38	505.5	30.4	317	1	US-08-463-090B-9	Sequence 9, Appl
39	505.5	30.4	317	2	US-08-874-347-18	Sequence 18, Appl
40	505.5	30.4	317	3	US-09-093-522-18	Sequence 18, Appl
41	505.5	30.4	527	4	US-09-739-455-2	Sequence 2, Appl
42	505	30.4	323	2	US-08-874-347-21	Sequence 21, Appl
43	505	30.4	323	3	US-09-093-522-21	Sequence 21, Appl
44	500.5	30.1	282	1	US-08-318-947A-19	Sequence 19, Appl
45	500.5	30.1	282	2	US-08-795-303-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-411-628-11
; Sequence 11, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-628-11

Query Match 71.2%; Score 1183.5; DB 4; Length 358;
Best Local Similarity 71.8%; Pred. No. 2.5e-116;
Matches 211; Conservative 41; Mismatches 41; Indels 1; Gaps 1;
QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVWKKIALREIRMLKQLKH 60
Db 2 MEKYEKIGKIGESYGVVFKCRNRTDQGVAVKVFLESEDDPVWKKIALREIRMLKQLKH 61
QY 61 PNLVNLIEVPRRRKQVHLVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCHTH 120
Db 62 PNLVNLLEVRKRRRLHLVFEYCDHTLVHLDRYQGVPEHLVKSVITWOTLQAVNFCCHK 121
QY 121 NCIHRDIKENILITKQGIKICDFGFAILL-PGDATDYVATRYRAPPELLVGDQTQYG 179
Db 122 NCIHRDVKNPILITKHSVVKICLDFGFAELLTGPSSDYTYDVATRYRSPPELLVGDQTQYG 181
QY 180 SSVDIWAIGCVFAELLTGPVPGKSDVDQLYIIRTLGKLIPRHOSIFKSNGFHGISI 239
Db 182 PVDVWAIGCVFAELLSGVPLWPGKSDVDQLYIRKTLGDLIPRHQOVFSTNQYFSGVKI 241
QY 240 PEPEDMETLEKPSDVHVPVLMFMKGLKQNPDDRLITCSQLLESSYFDSFQBAQ 293
Db 242 PDPEDMEPLEKFPNISYPALGLXGLHMDPTERTLCBQLLHHPYFENIREIE 295
RESULT 2
US-09-411-628-4
; Sequence 4, Application US/09411628
; Patent No. 6428994

; GENERAL INFORMATION:
 ; APPLICANT: University of Southern California
 ; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
 ; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
 ; FILE REFERENCE: 13761-707
 ; CURRENT APPLICATION NUMBER: US/09/411,628
 ; CURRENT FILING DATE: 1999-10-01
 ; EARLIER APPLICATION NUMBER: US 60/102,906
 ; EARLIER FILING DATE: 1998-10-02
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 566
 ; TYPE: PRT
 ; ORGANISM: Oryctolagus cuniculus
 ; US-09-411-628-4

Query Match 58.2%; Score 967.5; DS 4; Length 566;
 Best Local Similarity 54.6%; Pred. No. 2.9e-93;
 Matches 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKALREIRMLKQLKH 60
 DB 1 MEKYENGLVGEISYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKALREIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKRKRKHILVFEYCDHTLLNELRPNFNGVADGVKISVLWQTLQALNFCIH 120
 DB 61 ENLVNLLEVCKKKRWLVFEFVDHTILDDLELFPNGLDQVQVOKYLFQING:GFCHSH 120
 QY 121 NCIRHDIKPNILITKQGIKICDFGFAQIL-IPGDATDVVATRWYRAPPELLVGDTCYG 179
 DB 121 NIHRDIPENILVSQSGVVKLCDFGFARTLAAGEVYTDVATRWYRAPPELLVGDTCYG 180
 QY 180 SSVDIWAGCVFAELLTQGLPWGKSDVDQYLIIRTLGKLPHQSFKSGFFHGISI 239
 DB 180 SSVDIWAGCVFAELLTQGLPWGKSDVDQYLIIRTLGKLPHQSFKSGFFHGISI 239
 QY 240 PEPEDMETLEKFSVHPVAVNFKMCKLMMPPDRLTCSQLLESSYF--DSF-----QEA 292
 DB 240 PEPEDMETLEKFSVHPVAVNFKMCKLMMPPDRLTCSQLLESSYF--DSF-----QEA 292
 QY 293 QIK--RKARN-----EGRNRRRQ 309
 DB 301 QIKVQKDARNVSLSKKSNRRKKEK 324

RESULT 3
 US-09-411-628-10
 ; Sequence 10, Application US/09411628
 ; Patent No. 6428994
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Southern California
 ; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
 ; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
 ; FILE REFERENCE: 13761-707
 ; CURRENT APPLICATION NUMBER: US/09/411,628
 ; CURRENT FILING DATE: 1999-10-01
 ; EARLIER APPLICATION NUMBER: US 60/102,906
 ; EARLIER FILING DATE: 1998-10-02
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 493
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-411-628-10

Query Match 56.6%; Score 941.5; DS 4; Length 493;
 Best Local Similarity 53.1%; Pred. No. 1.3e-90;
 Matches 172; Conservative 68; Mismatches 69; Indels 15; Gaps 5;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKALREIRMLKQLKH 60

DB 1 MEKYENGLVGEISYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKALREIRMLKQLKH 60
 QY 61 PNLVNLIEVFRKRKRKHILVFEYCDHTLLNELRPNFNGVADGVKISVLWQTLQALNFCIH 120
 DB 61 ENLVNLLEVCKKKRWLVFEFVDHTILDDLELFPNGLDQVQVOKYLFQING:GFCHSH 120
 QY 121 NCIRHDIKPNILITKQGIKICDFGFAQIL-IPGDATDVVATRWYRAPPELLVGDTCYG 179
 DB 121 NIHRDIPENILVSQSGVVKLCDFGFARTLAAGEVYTDVATRWYRAPPELLVGDTCYG 180
 QY 180 SSVDIWAGCVFAELLTQGLPWGKSDVDQYLIIRTLGKLPHQSFKSGFFHGISI 239
 DB 180 SSVDIWAGCVFAELLTQGLPWGKSDVDQYLIIRTLGKLPHQSFKSGFFHGISI 239
 QY 240 PEPEDMETLEKFSVHPVAVNFKMCKLMMPPDRLTCSQLLESSYF--DSF-----QEA 292
 DB 240 PEPEDMETLEKFSVHPVAVNFKMCKLMMPPDRLTCSQLLESSYF--DSF-----QEA 292
 QY 293 QIK--RKARN-----EGRNRRRQ 309
 DB 301 QIKVQKDARNVSLSKKSNRRKKEK 324

RESULT 4
 US-08-874-347-25
 ; Sequence 25, Application US/08874347
 ; Patent No. 5863741
 ; GENERAL INFORMATION:
 ; APPLICANT: Linper, Andrew H.
 ; APPLICANT: Leof, Edward B.
 ; APPLICANT: Thomas, Charles F.
 ; APPLICANT: Gustafson, Michael P.
 ; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
 ; TITLE OF INVENTION: CARINII
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C., P.A.
 ; STREET: 60 South Sixth Street, Suite 3300
 ; CITY: Minneapolis
 ; STATE: MN
 ; COUNTRY: USA
 ; ZIP: 55402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/874,347
 ; FILING DATE: 13-JUN-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ellinger, Mark S.
 ; REGISTRATION NUMBER: 34,812
 ; REFERENCE/DOCKET NUMBER: 07039/055001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612-335-5070
 ; TELEFAX: 612-288-9696
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 298 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-874-347-25

Query Match 34.3%; Score 570; DB 2; Length 298;
 Best Local Similarity 40.0%; Pred. No. 8.5e-52;
 Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGGVAVKKFVESEDDPVVKKALREIRMLKQLKH 60

[illegible]

RESULT 6
US-09-093-522-25
; Sequence 25, Application US/09093522
; Patent No. 6015700

GENERAL INFORMATION:
APPLICANT: Limper, Andrew H.
APPLICANT: Leof, Edward B.
APPLICANT: Thomas, Charles F.
APPLICANT: Gustafson, Michael P.
TITLE OF INVENTION: CPC2 PROTEIN KINASE FROM PNEUMOCOCCUS
TITLE OF INVENTION: CARINI
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C., P.A.
STREET: 60 South Sixth Street, Suite 3300
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatib.e
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/093,522
FILING DATE: 08-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/874,347
FILING DATE: 13-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PS-09-093-522-25

	Query Match	34.3%	Score 570;	DB 3;	Length 298;
	Best Local Similarity	40.0%	Prod. No. 8.5e-52;		
	Matches 118;	Conservative 60;	Mismatches 99;	Indels 18;	Gaps 6
QY	1	MEKYEKLAKTGGESYGVVFKCRNKTSGOVAVKVKFVESEDDPVVKIALRETEMLKQLKH	60		
Db	1	MENFQKVEKGEGTGVVYKARNKLTGCVVALKKIRLDTETEGVPSTAREISLLKELNH	60		
QY	61	PNLVNLVIEVFRKRKQHLVPEYCDHTLLNELERNP-NGVADGVYKSVLMQTIQALNFCHI	119		


```
; SEQ ID NO 113
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CDK2-EGFP fusion
US-09-417-197-113

Query Match          34.3%; Score 570; DB 4; Length 544;
Best Local Similarity 40.0%; Pred. No. 2.1e-51;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MENFQVKEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPTAIREISLLKELNH 60
QY 61 PNVLNLIIEVRRKRMHLVPEYCDHTLLNELEENP-NGVADGVKSVLMQTLQALNFCHI 119
DB 61 PNIVKLLDVHTENKLYLVEFLHQDLKKFMDASALTG:PLP:IKSY:FLQLQGLAFCHS 120
QY 120 HNCIHRDKPENILITKQGIKICDFGFAQIL-IPGDAYTVVATRWYRAPELLVGDTOY 178
DB 121 HRVLRDLKPQNLLINTEGAIKLADFLARAFGVPVRTYTHEVVTLMWYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQVLYLITRTLGKLIPIRHQSIFKSNGFHGI- 237
DB 181 STAVDIWSLGCIFAEMVTRRALPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY 238 SIPE--PEDMETLEEKFSVDHPV----ALNFMKGCLKMNPDDRLTCSQLESSEYF 286
DB 232 SMPDYKPSFFPKWARQDFSKVVPPLDEGSRLLSQMLHYDENKKAISAKAALAHPPF 286

RESULT 10
US-09-417-197-115
; Sequence 115, Application US/09417197
; Patent No. 6518221
; GENERAL INFORMATION:
; APPLICANT: Ole Thastrup, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative information Relating To An
; FILE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 115
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-CDK2 fusion
US-09-417-197-115

Query Match          34.3%; Score 570; DB 4; Length 544;
Best Local Similarity 40.0%; Pred. No. 2.1e-51;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 247 MENFQVKEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPTAIREISLLKELNH 306
QY 61 PNVLNLIIEVRRKRMHLVPEYCDHTLLNELEENP-NGVADGVKSVLMQTLQALNFCHI 119
DB 307 PNIVKLLDVHTENKLYLVEFLHQDLKKFMDASALTG:PLP:IKSY:FLQLQGLAFCHS 366
QY 120 HNCIHRDKPENILITKQGIKICDFGFAQIL-IPGDAYTVVATRWYRAPELLVGDTOY 178
DB 367 HRVLRDLKPQNLLINTEGAIKLADFLARAFGVPVRTYTHEVVTLMWYRAPELLGSKYY 426
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQVLYLITRTLGKLIPIRHQSIFKSNGFHGI- 237
DB 427 STAVDIWSLGCIFAEMVTRRALPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 477

; SEQ ID NO 113
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CDK2-EGFP fusion
US-09-417-197-113

Query Match          34.3%; Score 570; DB 4; Length 544;
Best Local Similarity 40.0%; Pred. No. 2.1e-51;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MENFQVKEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPTAIREISLLKELNH 60
QY 61 PNVLNLIIEVRRKRMHLVPEYCDHTLLNELEENP-NGVADGVKSVLMQTLQALNFCHI 119
DB 61 PNIVKLLDVHTENKLYLVEFLHQDLKKFMDASALTG:PLP:IKSY:FLQLQGLAFCHS 120
QY 120 HNCIHRDKPENILITKQGIKICDFGFAQIL-IPGDAYTVVATRWYRAPELLVGDTOY 178
DB 121 HRVLRDLKPQNLLINTEGAIKLADFLARAFGVPVRTYTHEVVTLMWYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQVLYLITRTLGKLIPIRHQSIFKSNGFHGI- 237
DB 181 STAVDIWSLGCIFAEMVTRRALPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY 238 SIPE--PEDMETLEEKFSVDHPV----ALNFMKGCLKMNPDDRLTCSQLESSEYF 286
DB 232 SMPDYKPSFFPKWARQDFSKVVPPLDEGSRLLSQMLHYDENKKAISAKAALAHPPF 286

RESULT 11
US-09-411-628-13
; Sequence 13, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER FILING DATE: 1998-10-06
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(298)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-411-628-13

Query Match          34.2%; Score 569; DB 4; Length 298;
Best Local Similarity 40.0%; Pred. No. 1.1e-51;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MENFQVKEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPTAIREISLLKELNH 60
QY 61 PNVLNLIIEVRRKRMHLVPEYCDHTLLNELEENP-NGVADGVKSVLMQTLQALNFCHI 119
DB 61 PNIVKLLDVHTENKLYLVEFLHQDLKKFMDASALTG:PLP:IKSY:FLQLQGLAFCHS 120
QY 120 HNCIHRDKPENILITKQGIKICDFGFAQIL-IPGDAYTVVATRWYRAPELLVGDTOY 178
DB 121 HRVLRDLKPQNLLINTEGAIKLADFLARAFGVPVRTYTHEVVTLMWYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQVLYLITRTLGKLIPIRHQSIFKSNGFHGI- 237
DB 181 STAVDIWSLGCIFAEMVTRRALPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY 238 SIPE--PEDMETLEEKFSVDHPV----ALNFMKGCLKMNPDDRLTCSQLESSEYF 286
DB 232 SMPDYKPSFFPKWARQDFSKVVPPLDEGSRLLSQMLHYDENKKAISAKAALAHPPF 286

RESULT 12
US-08-074-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
```

COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,347
FILING DATE: 13-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055001
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 294 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-874-347-26

Query Match 34.2%; Score 568.5; DB 2; Length 294;
Best Local Similarity 37.6%; Pred. No. 1.2e-51;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYELAKTGGSGYGVFKCRNKTSQGVAVKVFSEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEQYEEKIGETGYGVYVYRDKVNETIALKIRLEQDEGVSTAIRISLLKEMHH 60
QY 61 PNLVNLIEVFRKRKHLYVFXCDHTLLN-----ELEENPNGVADGVKSVLWOTLOAL 114
DB 61 GNIVRLHDVHSEKRIYLYVFXLDLKKFMSCEFAKNT-----LKSXYLIQLRGV 115
QY 115 NFCHHNCIHRDIKPNILITKO-GIIKICDFGFAQIL-IPGDVATVYVATRWYRAPELL 172
DB 116 AYCHSHRVLRLDKPQNLIDRRTNALKLADFLARAFGIPVRTTHTHEVTLWYRAPEIL 175
QY 173 VGTQYGVSDIWAICVFAELLTGQPLWPKGSDVDQLYLIIRTLGKLI PRHOSIFKNSG 232
DB 176 LGSRYSTPDMWMSVGCIPAEVYNQKPLFGDSEIDELFKIPRVLG--TPNEQS----- 227
QY 233 FFHGI-SIP-----EPEDMETLEKFSVHPVALNFMKGCMLKMPDRLTCSOLLE 282
DB 228 -WPGVSSLDPYKSAPFKWQAQLATI---VPTLDPAGLDLLSKMLRYEPNKRITARQALE 283
QY 283 SSYFDSFQEAQ 293
DB 284 HEYFKOLEMVQ 294

RESULT 13
US-09-093-522-26
Sequence 26, Application US/09093522
Patent No. 6015700
GENERAL INFORMATION:
APPLICANT: Linper, Andrew H.
APPLICANT: Leof, Edward B. F.
APPLICANT: Thomas, Charles F.
APPLICANT: Gustafson, Michael P.
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
TITLE OF INVENTION: CARINI
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C., P.A.
STREET: 60 South Sixth Street, Suite 3300
CITY: Minneapolis

STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/093,522
FILING DATE: 08-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/874,347
FILING DATE: 13-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055002
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 294 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-093-522-26

Query Match 34.2%; Score 568.5; DB 3; Length 294;
Best Local Similarity 37.6%; Pred. No. 1.2e-51;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYELAKTGGSGYGVFKCRNKTSQGVAVKVFSEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEQYEEKIGETGYGVYVYRDKVNETIALKIRLEQDEGVSTAIRISLLKEMHH 60
QY 61 PNLVNLIEVFRKRKHLYVFXCDHTLLN-----ELEENPNGVADGVKSVLWOTLOAL 114
DB 61 GNIVRLHDVHSEKRIYLYVFXLDLKKFMSCEFAKNT-----LKSXYLIQLRGV 115
QY 115 NFCHHNCIHRDIKPNILITKO-GIIKICDFGFAQIL-IPGDVATVYVATRWYRAPELL 172
DB 116 AYCHSHRVLRLDKPQNLIDRRTNALKLADFLARAFGIPVRTTHTHEVTLWYRAPEIL 175
QY 173 VGTQYGVSDIWAICVFAELLTGQPLWPKGSDVDQLYLIIRTLGKLI PRHOSIFKNSG 232
DB 176 LGSRYSTPDMWMSVGCIPAEVYNQKPLFGDSEIDELFKIPRVLG--TPNEQS----- 227
QY 233 FFHGI-SIP-----EPEDMETLEKFSVHPVALNFMKGCMLKMPDRLTCSOLLE 282
DB 228 -WPGVSSLDPYKSAPFKWQAQLATI---VPTLDPAGLDLLSKMLRYEPNKRITARQALE 283
QY 283 SSYFDSFQEAQ 293
DB 284 HEYFKOLEMVQ 294

RESULT 14
US-09-457-040B-29
Sequence 29, Application US/09457040B
Patent No. 6387641
GENERAL INFORMATION:
APPLICANT: Vertex Pharmaceuticals Incorporated
APPLICANT: Beilon, Steve
TITLE OF INVENTION: Crystallized P38 Complexes
FILE REFERENCE: VPI/98-14
CURRENT APPLICATION NUMBER: US/09/457,040B
CURRENT FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 41

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:16:37 ; Search time 13.3233 Seconds
(without alignments)
2273.702 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1.662

Sequence: 1 MEKYEKLAKTSGSYGVVFK.....RKARNEGRNRQRQVLPLKS 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR 76:**

1: Pirt:**

2: Pirt:**

3: Pirt:**

4: Pirt:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183.5	71.2	358	2 S23383	protein kinase (EC
2	1175.5	70.7	376	2 S22745	serine/threonine k
3	581	35.0	305	2 S23382	protein kinase (EC
4	574.5	34.6	292	2 S40021	protein kinase (EC
5	571.5	34.4	294	1 A40444	protein kinase (EC
6	571.5	34.4	294	2 B40444	protein kinase (EC
7	570.5	34.3	302	1 I50474	protein kinase (EC
8	568.5	34.2	294	2 S22440	protein kinase (EC
9	567	34.1	298	2 A41227	protein kinase (EC
10	565.5	34.0	297	2 A37871	protein kinase (EC
11	561.5	33.8	294	2 S23095	protein kinase (EC
12	559	33.6	302	1 A44349	protein kinase (EC
13	557	33.5	297	1 S24913	protein kinase (EC
14	556.5	33.5	294	2 T49271	CELL DIVISION CON
15	556	33.5	297	2 A29539	protein kinase (EC
16	556	33.5	297	2 I45377	cyclin-dependent k
17	556	33.5	302	2 B44349	protein kinase (EC
18	554	33.3	301	1 S19209	protein kinase (EC
19	553	33.3	298	1 A44878	protein kinase (EC
20	551.5	33.2	311	2 S36619	protein kinase (EC
21	551	33.2	288	1 S42566	protein kinase (EC
22	550.5	33.1	294	1 JQ2243	protein kinase (EC
23	550.5	33.1	294	1 S57928	protein kinase (EC
24	550.5	33.1	294	1 S42049	protein kinase (EC
25	547.5	32.9	303	1 S06011	protein kinase (EC
26	547.5	32.9	308	1 S53538	protein kinase cdc
27	546.5	32.9	297	1 S12009	protein kinase (EC
28	545	32.8	292	1 A46365	protein kinase (EC
29	544.5	32.8	291	2 S23386	protein kinase (EC

30	544	32.7	233	2 JE0374	cyclin-dependent k
31	543	32.7	232	1 I49592	protein kinase (EC
32	543	32.7	232	1 A45091	protein kinase (EC
33	543	32.7	237	1 A36074	protein kinase (EC
34	543	32.7	301	1 S42101	protein kinase (EC
35	541.5	32.6	294	2 S23332	protein kinase (EC
36	540	32.5	296	2 S24386	protein kinase (EC
37	539	32.4	302	2 OKB785	protein kinase PHO
38	535.5	32.2	302	2 T17115	protein kinase cdc
39	530.5	31.9	346	2 A54820	CDK-activating pro
40	530	31.9	294	2 S51008	protein kinase (EC
41	528	31.8	232	2 S22441	protein kinase (EC
42	527.5	31.7	346	2 A56231	MO15/CDK-activatin
43	527	31.7	346	1 I78840	protein kinase (EC
44	525	31.6	332	1 S41003	protein kinase (EC
45	525	31.6	346	1 I48157	protein kinase (EC

ALIGNMENTS

RESULT 1

S23383

protein kinase (EC 2.7.1.37; cdc2-related KKIALRE - human

C:Species: Homo sapiens (man)

C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999

C:Accession: S23383; S22744

R:Mayerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; T

EMBO J. 11, 2909-2917, 1992

A:Title: A family of human cdc2-related protein kinases.

A:Reference number: S23392; MUID:92347325; PMID:1639063

A:Accession: S23383

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-358 <MEY>

A:Cross-references: EMBL:X66358; NID:g36614; PIDN:CAA47002.1; PID:g36615

C:Superfamily: Kinase-related transforming protein; protein kinase homology

C:Keywords: Atp; phosphotransferase; serine/threonine-specific protein kinase

F:3-278/Domain: protein kinase homology <Kin>

F:11-19/Region: protein kinase ATP-binding motif

F:34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match	71.2%	Score 1183.5;	DB 2;	Length 358;
Best Local Similarity	71.8%	Pred. No. 3.9e-49;		
Matches 211;	Conservative 41;	Mismatches 41;	Indels 1;	Gaps 1;
QY	1	MEKYEKLAKTSGSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH	60	
DB	2	MEKYEKIGKIGESYGVVFKCRNRTSGVAVKKFLESEDDPVVKIALREIRMLKQLKH	61	
QY	61	PNLVNLTVEFRKKQHLVFEYCDHTLLALELRNPNGVADGVKISVLWOTLOALNFECHT	120	
DB	62	PNLVNLTVEFRKKRRLHLVFEYCDHTLVHLDYQGVPEHLVKISITWOTLOANFECHK	121	
QY	121	NCIHRDKPENILITQGIKICDFGFAQLI- PGDAYTYVATRYRAPPELLVGTQYG	179	
DB	122	NCIHRDKPENILITKHSVVKLCDFGFARLLTGPSDYTYDYVATRYRSPPELLVGTQYG	181	
QY	180	SSVDYWAIGCVFAELITGQPLWFGKSDVDQLYLITLGLKLIPIRHOSIFKSGFFHGISI	239	
DB	182	PPVDVWAIGCVFAELISGVPFWFGKSDVDQLYLITLGLLPIRHQQVFTSQYFSGVKI	241	
QY	240	PEPEMETLEKFSVDVHPVALNFMKGLKXNPPDRLTCSQLLESSYFDSFQEAQ	293	
DB	242	PDPEMEPLEKFPNISYPALGLLKGCLHMDPTERTLTCEQLLHHPYFENIREIE	295	

RESULT 2

S22745

serine/threonine protein kinase KKIALRE (EC 2.7.1.-) - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Dec-1997

C:Accession: S22745


```
Db 120 HCRVRLHRELKLPQNLIDNKGVIKLADFLGARFGVGVVYVTHEVVLWYRAPSVLLGAS 179
QY 177 QYGVSSVDIWAIGCVFAELLTGPLWPGKSDVQVLYLRTLG-----KLIPIHQSS 226
Db 180 RYSTPDVMSIGTFEALATKAPLPHGSDIEQLFRIFCTGLGTPRNEVWPVESLPDYKN 239
QY 227 IF---KSNQFFHGSIPEDMETLEEKFSVDHPVALNFMKCKKXNPDRLTCSLLES 283
Db 240 TFRKXWSSG-----NIASTVKNLDKXGIDLTAKMLLYDPPKISARQAXTH 284
QY 284 SYFDSFOEAO: 294
Db 285 PYFCDLDKSTL 295

RESULT 8
S22440
protein kinase [EC 2.7.1.37] cdc2 homolog 1 - rice
C:Species: Oryza sativa (rice)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S22440
R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuki, I.; Utsugi,
Yol: Gen. Genet. 233, 10-16, 1992
A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory
A:Reference number: S22440; MUID:92293101; PMID:1376401
A:Accession: S22440
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1294 <HAS>
A:Cross-references: EMBL:X60374; NID:G20342; PIDN:CAA42922.1; PID:920343
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:15/Binding site: phosphate (Thr) (covalent) #status predicted
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.2%; Score 568.5; DB 2; Length 294;
Best Local Similarity 37.6%; Pred. No. 2.1e-20;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKLATGEGSYGVFKCRNKTSGQVAVKVFSEDDPVVKIALREIRMLKQLKH 60
Db 1 MEKYEKEKIGESTYGVVTRADKVTNETIALKRLGECGCVSTAIRISLSEKWH 60
QY 61 PNLVNLIIEVRRKRMHLYFEYCDHTLN-----ELERNPGVADGVKSVLWQTLQAL 114
Db 61 GNIVRLHDVTHSEKRIYLVFEYLDLKLKFMDSCEFAKNPT-----LIKSLYQILRGV 115
QY 115 NFCHINCHRDIKPENILCTTKQ-GIKICDFGPAQL-IPGDAYDYVATWYRAPELL 172
Db 116 AYCHSRVLRDLKPNLLIDRTNALKLADFLGARAFG:PVRTTTHEVTLWYRAPELL 175
QY 173 VGTQYGVSSVDIWAIGCVFAELLTGQPLWPGKSDVQVLYLRTLGKLIPIHQSSIFKSG 232
Db 176 LGSQYSTPDVMSVGCIFAEVMNQKPLPFGDSEIDCE:FKIPIRLG--TPNEQS----- 227
QY 233 FPHGI-SIP-----EPEDMETLEEKFSVDHPVALNFMKCKKXNPDRLTCSLLE 262
Db 228 -WPGVSSLPDYKSAFPWQAQCLATI---VPTLDPAGLLSKMLREYENKRIATARCALE 263
QY 283 SSYFDSFOEAO: 293
Db 284 HEYFKOLEMWQ 294

RESULT 9
A41227
protein kinase [EC 2.7.1.37] cdk2 - human
N:Alternate names: Egl homolog; protein kinase p34
C:Species: Homo sapiens (man)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
C:Accession: A41227; S16520
```

```
R:Ninomiya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A:Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a
A:Reference number: A41227; MUID:92020980; PMID:1717994
A:Accession: A41227
A:Molecule type: mRNA
A:Residues: 1-298 <NIN>
A:Cross-references: GB:M68520; NID:gl80177; PIDN:AAA35667.1; PID:gl80178
R:Tsai, L.H.; Harlow, E.; Meyerson, M.
Nature 353, 174-177, 1991
A:Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1F
A:Reference number: S17873; MUID:91367262; PMID:1653904
A:Accession: S17873
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-298 <TSA>
A:Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803
R:Eledge, S.J.; Spottswood, M.R.
EMBO J. 10, 2653-2659, 1991
A:Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc2
A:Reference number: S16520; MUID:91330891; PMID:1714386
A:Accession: S16520
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-176, 'S', 178-298 <ELL>
A:Cross-references: EMBL:X61622; NID:G29848; PIDN:CAA43807.1; PID:G29849
C:Genetics:
A:Gene: GDB:CDK2
A:Cross-references: GDB:128984; OMIM:116953
A:Map position: 12q13-12q13
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serir
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:14,160/Binding site: phosphate (Thr) (covalent) #status predicted
F:15/Binding site: phosphate (Tyr) (covalent) #status predicted
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.1%; Score 567; DB 2; Length 298;
Best Local Similarity 40.0%; Pred. No. 2.5e-20;
Matches 119; Conservative 60; Mismatches 99; Indels 19; Gaps 6;

QY 1 MEKYEKLATGEGSYGVFKCRNKTSGQVAVKVFSEDDPVVKIALREIRMLKQLKH 60
Db 1 MEKYEKEKIGESTYGVVTRADKVTNETIALKRLGECGCVSTAIRISLSEKWH 60
QY 61 PNLVNLIIEVRRKRMHLYFEYCDHTLNELERNP-NGVADGVKSVLWQTLQALNFC 119
Db 61 PNIVKLLDVHTENKLYLVFEFHQDLKFKMDASALTGIPLIKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGPAQL-IPGDAYDYVATWYRAPELLVGDYQY 178
Db 121 HRYLHRLDKPNLLINTEGAIKLADFLGARAFG:PVRTTTHEVTLWYRAPELLVGC 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVQVLYLRTLGKLIPIHQSSIFKSGFPHGI- 237
Db 181 STAVQVMSVGCIFAEVMNQKPLPFGDSEIDQIFRITLGT-----PDEVVWPGVT 231
QY 238 SIFE--PEMETLEEKFSVDHPV----ALNFMKCKKXNPDRLTCSLLESSYF 286
Db 232 SMPDYKPSFPKWARQDFSKVPPPLDDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 10
A37871
protein kinase [EC 2.7.1.37] cdk2 - African clawed frog
N:Alternate names: cell division control protein CDC2 homolog Egl
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 18-Jun-1999
C:Accession: A37871; S15866; I51662; S14410
R:Paris, J.; Le Guellec, R.; Couturier, A.; Le Guellec, K.; Omilli, F.; Camonis, J.; Me
Proc. Natl. Acad. Sci. U.S.A. 88, 1039-1043, 1991
A:Title: Cloning by differential screening of a Xenopus cDNA coding for a protein high
```


A:Reference number: A37871; MUID:91126051; PMID:1704128
A:Accession: A37871
A:Molecule type: mRNA
A:Residues: 1-297 <PAR>
A:Cross-references: GB:X14227
R:Le Guellec, R.
Submitted to the EMBL Data Library, January 1989
A:Reference number: S15866
A:Accession: S15866
A:Molecule type: mRNA
A:Residues: 1-92, R, 94-297 <DEA>
A:Cross-references: EMBL:X14227; NID:G64665; PIDN:CAA32443.1; ZID:364666
R:Olive, M.; There, N.; Philippe, M.; Le Pennec, J.P.; Lerivray, R.
Gene 151, 81-88, 1994
A:Title: Cloning of the Xenopus laevis cdk2 promoter and functional analysis in oocytes
A:Reference number: S15622; MUID:95129896; PMID:7828909
A:Accession: S15622
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-39 <OLI>
A:Cross-references: EMBL:U07979; NID:G473584; PIDN:AAA91213.1; PID:G473585
C:Genetics:
A:Gene: cdk2
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-25/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.0%; Score 565.5; DB 2; Length 297;
Best Local Similarity 37.8%; Pred. No. 3e-20;
Matches 116; Conservative 65; Mismatches 85; Indels 4; Gaps 6;

QY 1 MEKYELAKTGGSGYGVFKCRNKTSGQVAVKVFESDDPVVKKIALREIRMLKQKH 60
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
1 MENFQKVEKIGGTGVVYKARNRETGEIVALKKRLDTETEGVSPSTAIRSLLKELNH 60

QY 61 PNLVNLIEVFRKKRKHVFEYCDHTLLNELRNPNGVAD-GVTKSVLWQTLQALNFCHI 119
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 119
61 PNLVNLIEVFRKKRKHVFEYCDHTLLNELRNPNGVAD-GVTKSVLWQTLQALNFCHI 119

QY 120 HNCIHRDKPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDVATWYRAPELLVGDTQY 178
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 178
121 HRVLRDLKPNLLIDRRNTSLKDLADFLARAFGIPVRTFTHEWTLWYRAPELLVGS 180

QY 179 GSSVDIWAICGVFAELLTCQPLWPKSDVDQYLIIRTLG-----KUIPRHQSI 228
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 228
181 STAVDIWSLGCIFAEMITRAJFPFGDSEIDQLFRPFTLTGTPDEVSWPGVITMPOYKSTF 240

QY 229 KSNGFHFGISIPEDMETLEEKFSVHPVALNFKMCKLKNPDDRLTCSQLLESYFDS 288
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 288
241 P-----KMIROFSKVP-----PDEGGRDLAQLQ---YDS 272

QY 289 FQEAQIK 295
DB :
272 NKRISAK 278

RESULT 11
S23095
protein kinase (EC 2.7.1.37) cdc2 - Arabidopsis thaliana
N:Alternate names: cdc2 protein homolog; CDC2a protein; cell division control protein 2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
R:Accession: S23095; A48984; JQ1337; JQ0967; S18202
R:Imajuku, Y.; Hirayama, T.; Endoh, H.; Oka, A.
FEBS Lett. 304, 73-77, 1992
A:Title: Exon-intron organization of the Arabidopsis thaliana protein kinase genes CDC2a
A:Reference number: S23095; MUID:92316202; PMID:1618302
A:Accession: S23095
A:Molecule type: DNA
A:Residues: 1-294 <IMA>
A:Cross-references: EMBL:D10850; NID:G217848; PIDN:BA01623.1; PID:G217849

R:Inze, D.; Ferreira, P.; Hemerly, A.; Van Montagu, M.
Biochem. Soc. Trans. 20, 80-84, 1992
A:Title: Control of cell division in plants.
A:Reference number: A48984; MUID:92335744; PMID:1634002
A:Accession: A48984
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-294 <INZ>
A:Experimental source: flower
R:Hirayama, T.; Imajuku, Y.; Anai, T.; Matsui, M.; Oka, A.
Gene 105, 159-165, 1991
A:Title: Identification of two cell-cycle-controlling cdc2 gene homologs in Arabidopsis
A:Reference number: JQ1337; MUID:92039027; PMID:1937013
A:Accession: JQ1337
A:Molecule type: mRNA
A:Residues: 1-294 <HIP>
A:Cross-references: EMBL:X57839; NID:G16218; PIDN:CAA40971.1; P:ID:G16219
R:Ferreira, P.C.G.; Hemerly, A.S.; Villarroel, R.; Van Montagu, M.; Inze, D.
Plant Cell 3, 531-540, 1991
A:Title: The Arabidopsis functional homolog of the p34cdc2 protein kinase.
A:Reference number: JQ0967; MUID:93005715; PMID:1840925
A:Accession: JQ0967
A:Molecule type: mRNA
A:Residues: 1-294 <FER>
A:Experimental source: flower
C:Comment: The protein is a key component of the eukaryotic cell cycle.
C:Genetics:
A:Gene: cdc2
A:Introns: 3/3; 65/2; 105/3; 163/3; 218/2; 247/3; 265/3
C:Function:
A:Description: phosphotransferase; protein kinase; required for G1 to S-phase transition
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-25/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 33.8%; Score 561.5; DB 2; Length 294;
Best Local Similarity 37.1%; Pred. No. 4.5e-20;
Matches 111; Conservative 70; Mismatches 93; Indels 25; Gaps 6;

QY 1 MEKYELAKTGGSGYGVFKCRNKTSGQVAVKVFESDDPVVKKIALREIRMLKQKH 60
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
1 MDQYKVEKIGGTGVVYKARNRETGEIVALKKRLDTETEGVSPSTAIRSLLKELNH 60

QY 61 PNLVNLIEVFRKKRKHVFEYCDHTLLNELRNPNGVAD-GVTKSVLWQTLQALNFCHI 119
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 119
61 SNIVKLDQVHSEKRLVFEYLDLKKHWDSTPDSKDLHMKITLYQLRGIAVCHS 123

QY 120 HNCIHRDKPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDVATWYRAPELLVGDTQY 177
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 177
121 HRVLRDLKPNLLIDRRNTSLKDLADFLARAFGIPVRTFTHEWTLWYRAPELLVGS 180

QY 178 YGSSVDIWAICGVFAELLTCQPLWPKSDVDQYLIIRTLG-----KUIPRHQSI 227
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 227
181 YSTPVDIWSVGCIFAEMISQKPLFGDSEIDQLFKIPRIMGTPYEDTWGVTSLPDKYSA 240

QY 228 FKSNGFHFGISIPEDMETLEEKFSVHPVALNFKMCKLKNPDDRLTCSQLLESYF 286
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 286
241 FPK-----WKPTDLETF---VPLDQGVLLSKMLMDPTKRNARAALEHYF 287

RESULT 12
A44349
protein kinase (EC 2.7.1.37) cdc2-A [similarity] - African clawed frog
N:Alternate names: maturation-promoting factor p34cdc2 chain A
C:Species: Xenopus laevis (African clawed frog)
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: A44349
R:Pickham, K.M.; Meyer, A.N.; Li, J.; Donoghue, D.J.
Mol. Cell. Biol. 12, 3192-3203, 1992

A:Title: Requirement of Mos(Xe) protein kinase for meiotic maturation of Xenopus oocytes
A:Reference number: A44349; PMID:92318937; PMID:137775

A:Accession: A44349
A:Molecule type: mRNA
A:Residues: 1-302 <PIC>
A:Cross-references: GB:M60680; NID:G214022; PIDN:AAA63561.1; PID:G214023
A:Experimental source: oocytes
A:Note: sequence extracted from NCBI backbone (NCBIN:107682, NCBI:P:07683)
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33.51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 33.6%; Score 559; DB 1; Length 302;
Best Local Similarity 36.7%; Pred. No. 6e-20;
Matches 117; Conservative 68; Mismatches 102; Indels 32; Gaps 6;
QY 1 MEKYEKLAKTGGSGYGVFKCNKTSQGVAVKVFSEDDPVVKKIALREIRMLKQKH 60
DB 1 MDEYTKIEIGEGTYGVVYKGRHATGQVAMKKIRJENESEGVFSTA-REISLKEIQH 60
QY 61 PNLVNLIEVFRKRRKXHLVFEYCDHTLLNELRNPNQ--VADGVKSVLMOTLCALAFCH 118
DB 61 PNYVCLLDVLMGDSRLYLIFELSLMDLKKYLDSPSGQYIDTMLVKSLYLO-LGIVFCH 120
QY 119 HNCIHRDIKPNILITKOGIIRKICDFGFAQL-IPGDYTDYVATRYRAPELLVGDTQ 177
DB 121 SRVLHRLDKPNLLIDSGVKIKADFGARAFGIPRVVTHEVVTLMYRAPEVLGSR 180
QY 178 YGSSVDIWAIGCVFAELTGQPLWPKGSDVDQLYLIIRTLGKLI-PRHQSI-IPKSN- 237
DB 181 YSTPVDWISGTFPAELATKPLFHGSEIDQLFRIFALG--TPNNE----- 226
QY 238 SIPEPEDMETLEEF-----SDVHPVALNFMKGLKXNPPDRLTCSQLLESSY 263
DB 227 VMPVESLQDYXNTPKWKPGSLASHVKNLDENGLLSKMLVYDPAPKRISGKALKHP 286
QY 286 FDSFOEAQIKR 296
DB 287 FDDLKSSL---PDNQIRN 302

RESULT 13

S24913

Protein kinase (EC 2.7.1.37) cdc2 [similarity] - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000

C:Accession: S24913

R:Kanaoka, Y.; Nojima, H.; Okayama, H.

submitted to the EMBL Data Library, July 1991

A:Description: Nucleotide sequences of cDNAs encoding rat cdc2 + and cyclin 2.

A:Reference number: S20658

A:Accession: S24913

A:Molecule type: mRNA

A:Residues: 1-297 <KAN>

A:Cross-references: EXBL:X60767; NID:G57533; PIDN:CAA43177.1; PID:G57534

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase

F:2-256/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase ATP-binding motif

F:33.51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 33.5%; Score 557; DB 1; Length 297;

Best Local Similarity 39.1%; Pred. No. 7.4e-20;

Matches 122; Conservative 63; Mismatches 99; Indels 32; Gaps 8;

QY 1 MEKYEKLAKTGGSGYGVFKCNKTSQGVAVKVFSEDDPVVKKIALREIRMLKQKH 59

DB 1 MEDYTKIEIGEGTYGVVYKGRHRTTQIVAMKKIRLESEEG-VPSTA-REISLKEIQH 59

QY 60 HNLVNLIEVFRKRRKXHLVFEYCDHTLLNELRNPNQ--VADGVKSVLMOTLCALAFCH 117

DB 60 HNLVNLIEVFRKRRKXHLVFEYCDHTLLNELRNPNQ--VADGVKSVLMOTLCALAFCH 119
QY 118 HNCIHRDIKPNILITKOGIIRKICDFGFAQL-IPGDYTDYVATRYRAPELLVGDT 176
DB 120 HRRVHRLDKPNLLIDDKGTIKLADFGARAFGIPRVVTHEVVTLMYRSEVLLGSA 179
QY 177 YGSSVDIWAIGCVFAELTGQPLWPKGSDVDQLYLIIRTLGKLI-PRHQSI-IPKSN- 236
DB 180 RYSTPVDWISGTFPAELATKPLFHGSEIDQLFRIFALG--TPNNE----- 226
QY 237 SIPEPEDMETLEEFSDVHPVA-----LNFMKGLKXNPPDRLTCSQLLESS 284
DB 227 VMPVESLQDYXNTPKWKPGSLASHVKNLDENGLLSKMLVYDPAPKRISGKALKHP 285
QY 285 YFDSFOEAQIKR 296
DB 286 YFDDL-DNQIKK 296

RESULT 14

T49271

CELL DIVISION CONTROL PROTEIN 2 HOMOLOG A - Arabidopsis thaliana

N:Alternate names: protein T21J18.20

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000

C:Accession: T49271

R:Rieger, M.; Gabe, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Newes, H.W.; Rudd, S.

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z25021

A:Accession: T49271

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-294 <RIE>

A:Cross-references: EMBL:AL132963; GSPDB:GN00061; ATSP:T21J18.20

A:Experimental source: cultivar Columbia; BAC clone T21J18

C:Genetics:

A:Gene: ATSP:T21J18.20

A:Map position: 3

A:Insertions: 3/3; 67/3; 105/3; 163/3; 218/2; 247/3; 265/3

C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 33.5%; Score 556.5; DB 2; Length 294;

Best Local Similarity 36.8%; Pred. No. 7.8e-20;

Matches 110; Conservative 70; Mismatches 94; Indels 25; Gaps 6;

QY 1 MEKYEKLAKTGGSGYGVFKCNKTSQGVAVKVFSEDDPVVKKIALREIRMLKQKH 60

DB 1 MQYEKVEKIGEGTYGVVYKARDKVTNETIALKIRLQDEGVPSTAIRISLKEHQH 60

QY 61 PNLVNLIEVFRKRRKXHLVFEYCDHTLLNELRNPNQVAD-GVIKSVLMOTLCALAFCH 119

DB 61 SNIVYDDVHSEKRLYLVEYLDLKKHMDSTPDFSKDLHMKTYDYLRLGIAYCHS 120

QY 120 HNCIHRDIKPNILITKOGIIRKICDFGFAQL-IPGDYTDYVATRYRAPELLVGDTQ 177

DB 121 HRVLHRLDKPNLLIDRRNTSLKADFGARAFGIPRVVTHEVVTLMYRAPELLGSHH 180

QY 178 YGSSVDIWAIGCVFAELTGQPLWPKGSDVDQLYLIIRTLG-----KLIPRHOSI 227

DB 181 YSTPVDWISGTFPAELATKPLFHGSEIDQLFRIFALG--TPNNE----- 240

QY 228 FKSNGFFHGISIPEDMETLEEFSDVHPVALNFMKGLKXNPPDRLTCSQLLESSY 286

DB 241 FPK-----WKPTDLET---VPNLDPDGVLLSKMLMDPTKRLNARAALHEHYF 287

RESULT 15

A29539

protein kinase (EC 2.7.1.37) cdc2 - human

N:Alternate names: cell division control protein 2 (CDC2)

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jul-2000

C:Accession: A29539

R;Lee, M.G.; Nurse, P.
Nature 327, 31-35, 1987
A:Title: Complementation used to clone a human homologue of the fission yeast cell cycle
A:Reference number: A29539; XJID:87201915; PMID:3553962
A:Accession: A29539
A:Molecule type: mRNA
A:Residues: 1-297 <LEE>
A:Cross-references: GB:X05360; NID:G29838; PIDN:CAA28963.1; PID:G29833
C:Genetics:
A:Gene: GDB:CDC2
A:Cross-references: GDB:119052; OXIM:116940
A:Map position: 10q21.1-10q21.1
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:14,161/Binding site: phosphate (Thr) (covalent) #status predicted
F:15/Binding site: phosphate (Tyr) (covalent) #status predicted
F:33,51,128,130/Active site: Lys, Gln, Asp, Lys #status predicted

Query Match 33.5% Score 556; DB 2; Length 297;
Best Local Similarity 39.4%; Pred. No. 8.3e-20;
Matches 123; Conservative 61; Mismatches 96; Indels 32; Gaps 8;

QY 1 MEKYELAKTGGSYGVVFCRNKTSQVAVVKKP-VESEDDPVYKKIALBEIRYKQIK 59
DB 1 MEDYTKIEIGEGTYGVVYKGRHKITGQVAVMKIRLESEEG-VFSTAIRISLKEIR 59

QY 60 HPNLVNLIEVFRKRKHVLFVEYCDHTLLNELEARNPG--VADGVIKSVLWQTLQALNFC 117
DB 60 HPNIVSLQDVLMDSDSLYLIFELSLMDLKKYLCISPPGQYMDSSLVSKSYLQILOGVFC 119

QY 118 HINCIHROIKPNILITKQGIKICDGFQAQIL-IPGDAYTDYVATRWYRAPELNGDT 176
DB 120 HSRVYLRDLKPCNLLIDPKTIKADFGARAFGPIRIVYTHEVYTLWYRSPEVLLQSA 179

QY 177 QYGSVDIWAIGCVFAELLTGQPLMPKSDVQJLYLIHNTLGKLPRQS-FKSKGFTFG 236
DB 180 RYS:PTVIMSIGTIFAEIATKPLFGDSEIQQLFRIFALG--TPNNE----- 226

QY 237 ISIFEPEDMETLEEKFSQVHPVA-----LNFYKGCCKXNPDRITCSQLJESS 284
DB 227 -VMFEVESLQDYKNTFPKPKGSLASHVKNLQDGLDLSKMLIYDPAKRISGKQALNHP 285

QY 285 YEDSPQEAQIKR 296
DB 286 YFNDL-DNQIKK 296

Search completed: November 13, 2003, 14:21:17
Job time: 14.3233 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:12:56 ; Search time 9.51662 Seconds
(without alignments)
1556.584 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1662

Sequence: 1 MEKYEKLAKTGGSGYGVFK.....RKARNEGRNRQQV..PLKS 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026725 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1183.5	71.2	358	1	Q00532 homo sapien
2	720.5	43.4	1030	1	STK1A_HUMAN
3	581	35.0	305	1	CDK5_HUMAN
4	574.5	34.6	292	1	CDK2H_DICDI
5	571.5	34.4	297	1	CDK2_XENLA
6	570.5	34.3	302	1	CDK2_CARAU
7	568.5	34.2	294	1	CC21_ORYSA
8	567.5	34.1	303	1	CDK2_ORYCA
9	567	34.1	298	1	CDK2_HUMAN
10	566.5	34.1	294	1	CDK2_MAIZE
11	563	33.9	298	1	CDK2_CRIGR
12	561.5	33.8	294	1	CDK2_ARATH
13	561.5	33.8	303	1	CDK2_ORYLA
14	561	33.8	298	1	CDK2_RAT
15	559.5	33.7	303	1	CDK2_ORYCU
16	559.5	33.7	303	1	CDK2_ORYLU
17	559	33.6	298	1	CDK2_MESAU
18	559	33.6	302	1	CDK2_XENLA
19	558.5	33.6	294	1	CDK2_CHERU
20	557	33.5	297	1	CDK2_RAT
21	556	33.5	297	1	CDK2_BOVIN
22	556	33.5	297	1	CDK2_HUMAN
23	556	33.5	302	1	CDK2_XENLA
24	554	33.3	301	1	CDK2_TRYBB
25	553	33.3	298	1	CDK2_CARAU
26	551.5	33.2	294	1	CDK2_VIGUN
27	551.5	33.2	311	1	CC23_TRYBB
28	551	33.2	288	1	CC23_PLAPFK
29	550.5	33.1	294	1	CDK2_VIGAC
30	548	33.0	297	1	CDK2_MOUSE
31	547.5	32.9	303	1	CDK2_CHICK
32	546.5	32.9	297	1	CDK2_DROME
33	545	32.8	292	1	CDK5_BOVIN

34	545	32.8	292	1	CDK5_RAT	Q03114 rattus norv
35	544	32.7	292	1	CDK5_HUMAN	Q00535 homo sapien
36	543	32.7	292	1	CDK5_MOUSE	P49615 mus musculu
37	543	32.7	301	1	CC21_TRYCO	P54664 trypanosoma
38	541.5	32.6	294	1	CC22_MEDSA	Q05006 medicago sa
39	540	32.5	296	1	CDK2_DICDI	P34112 dictyosteli
40	536	32.3	292	1	CDK5_XENLA	P51166 xenopus lae
41	535.5	32.2	294	1	CC2A_ANTMA	Q08772 antirrhinum
42	533	32.1	302	1	CDK2_RANDY	Q9W739 rana dybows
43	533	32.1	305	1	PH85_YEAST	P17157 saccharomyc
44	530.5	31.9	346	1	CDK7_HUMAN	P50613 homo sapien
45	523	31.8	294	1	CDK5_DROME	P48609 drosophila

ALIGNMENTS

RESULT 1						
KK1A_HUMAN						
ID	KK1A_HUMAN	STANDARD;	PRT;	358 AA.		
AC	Q00532;					
DT	01-APR-1993 (Rel. 25, Created)					
DT	01-APR-1993 (Rel. 25, Last sequence update)					
DT	16-OCT-2001 (Rel. 40, Last annotation update)					
DE	Serine/threonine-protein kinase KK1AURE (EC 2.7.1.1) (Cyclin-dependent					
DE	kinase-like 1).					
GN	CDKL1.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RA	MEDLINE=92347325; PubMed=1639063;					
RA	Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,					
RA	Harlow E., Tsai L.-H.;					
RT	"A family of human cdc2-related protein kinases.";					
RL	EMBO J. 11:2909-2917(1992).					
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.					
CC	CDK2/CDXA SUBFAMILY.					

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X66356; CAA47002.1; -
EMBL; X66359; CAA47002.1; JOINED.
PIR; S23383; S23383.
HSSP; P24941; HCL.
Genew; HGNC:1781; CDKL1.
GK; Q00532; -
MIM; 603441; -
GO; GO:0004693; F:cyclin-dependent protein kinase activity; TAS.
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
GO; GO:0000074; P:regulation of cell cycle; TAS.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SMC0220; S_TKC; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
Transferase; Serine/threonine-protein kinase; ATP-binding.
DOMAIN 5 288 PROTEIN_KINASE.
FT NP_BIND 11 19 ATP (BY SIMILARITY).
FT BINDING 34 34 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
SQ SEQUENCE 358 AA; 41834 MW; 88344321F24B77C6 CRC64;

```
Query Match 71.2%; Score 1183.5; DB 1; Length 358;
Best Local Similarity 71.8%; Pred. No. 9.1e-82;
Matches 211; Conservative 41; Mismatches 4; Indels 1; Gaps 1;

QY 1 MEKYEKLAKTEGGSGVGVFKCRNKTSGGWAIVKVFSESDPVPVKKIALREIRMLKQLKH 60
DB 2 MEKYEKIGKIGEGSGVGVFKCRNRDTGQIVAKFLESESDPVPVKKIALREIRMLKQLKH 61
QY 61 PNLVNLIEVFRKRKHVLVEFYCOHTLNE:LERPNMGVAGGVKSVLWQTLQALNFCIH 120
DB 62 PNLVNLLEVFRKRKHVLVEFYCOHTLNE:DRYQGVPEHLVKSITWQTLQAVNFCIKH 122
QY 121 NCIHRRDIPENILITKQGIKICDFGAQILI-PCDAYTDVATRWYRAPPELLVGDQYQ 179
DB 122 NCIHRRDIPENILITKHSVIKLCDFGARLLTGSQDYTDVATRWYRAPPELLVGDQYQ 181
QY 180 SSVDTWAIAGCVFAELLTGQPLWPGKSDVDQVLYLIRTLGKLIIPHQSIFKSNFFHGISI 239
DB 182 PFDVWAIAGCVFAELLTSGVPLWPGKSDVDQVLYLIRTKLTGLIIPHQGVFSTNQYFSGVKI 241
QY 240 PEPDMETLEEKFSQVHPVALNFMKGCILKMPDRLTCSQLLESSYFDSFQEAQ 293
DB 242 PDPEDMELELKFNPNISYPALGLLGLGCLHMDPTERTCEQJHHYPENIREIE 295

RESULT 2
STK9 HUMAN STANDARD; PRT: 1C30 AA.
ID STK9 HUMAN STANDARD; PRT: 1C30 AA.
AC 076039; Q4198;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein kinase 9 (EC 2.7.1.37).
GN STK9.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9838628; PubMed=9721213;
RA Montini E., Andolfi G., Catuso A., Buchner G., Walpole S.M.,
RA Mariani M., Consalez G.G., Trump D., Balabio A., Franco B.;
RT "Identification and characterization of a novel serine-threonine
RT kinase gene from the Xp22 region.";
RL Genomics 51:427-433(1998).
RN [2]
RP SEQUENCE OF 339-789 FROM N.A.
RA Krause S.W., Rehl M., Kreutz M., Schwarzfischer J., Paulauskis J.D.,
RA Andreessen J.D.;
RT "Differential screening leads to novel genetic markers of monocyte
RT to macrophage maturation.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- ENZYME REGULATION: COULD BE ACTIVATED BY PHOSPHORYLATION ON
CC TYROSINE AND THREONINE. PROBABLY BOTH PHOSPHORYLATIONS ARE
CC REQUIRED FOR ACTIVITY.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, KIDNEY, PROSTATE,
CC OVARY, PLACENTA, PANCREAS AND TESTIS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-1C IS THE INITIATOR.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 415.
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC -----
DR EMBL; Y15057; CRA75342.1;
DR EMBL; X89059; CAA61445.1; ALT_FRAME.
DR HSSP; P24941; IB38.
DR Genew; HGNC:11411; STK9.
DR GK; O76039;
DR MIM; 300203;
DR GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 13 297 PROTEIN KINASE.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 135 135 BY SIMILARITY.
FT MOD_RES 169 169 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 171 171 (POTENTIAL).
FT MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 171 171 (POTENTIAL).
FT DOMAIN 784 789 POLY-LYS.
FT CONFLICT 339 340 HR -> GT (IN REF. 2).
FT CONFLICT 541 541 L -> W (IN REF. 2).
FT CONFLICT 731 764 MISSING (IN REF. 2).
SQ SEQUENCE 1030 AA; 115537 MW; 8A1C9C438610EF08 CRC64;
Query Match 43.4%; Score 720.5; DB 1; Length 1030;
Best Local Similarity 44.4%; Pred. No. 1.6e-46;
Matches 129; Conservative 69; Mismatches 99; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTEGGSGVGVFKCRNKTSGGWAIVKVFSESDPVPVKKIALREIRMLKQLKH 60
DB 10 MNKFELGVGVGEGAYGVVLCRHKHETHEIVAIFKFKDSENEEYKETTLEKMLRLTKQ 69
QY 61 PNLVNLIEVFRKRKHVLVEFYCOHTLNE:LERPNMGVAGGVKSVLWQTLQALNFCIH 120
DB 70 ENIVELKAFRRRGKLYLVFEVVEKQMLELEENPVGVPPEKVSYYIQL:KAIPWCHQN 129
QY 121 NCIHRRDIPENILITKQGIKICDFGAQILI-PCDAYTDVATRWYRAPPELLVGDQYQ 178
DB 130 DIVHRDIPENILISHNDVLCDFGARLLTGSQDYTDVATRWYRAPPELLVGDQYQ 188
QY 179 GSSVDWAIAGCVFAELLTGQPLWPGKSDVDQVLYLIRTLGKLIIPHQSIFKSNFFHGISI 238
DB 189 GKSVDWMSVGCILGELSDGQPLFGESIDQFTIQVGLPGLPSEOMKLFYSNPFHGLR 248
QY 239 IPEPEDMETLEEKFSQVHPVALNFMKGCILKMPDRLTCSQLLESSYFDSFQEAQIKR 297
DB 249 FPAVNHQPSLERRYLGILKSVLLDKMKNLLKLPADRYLTQCLN---HPTFQ-CRLDLR 305
QY 298 ARNEGNRRRQOV 310
DB 306 SPSRSARRRPYHV 318

RESULT 3
CDK3_HUMAN STANDARD; PRT: 305 AA.
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division protein kinase 3 (EC 2.7.1.-).
GN CDK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=92347325; PubMed=1639063;
 RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorfa C., Nelson C.,
 RA Harlow E., Tsai L.-H.;
 RT "A family of human cdc2-related protein kinases.";
 RL EMBO J. 11:2909-2917(1992).
 CC -|- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
 CC INTERACTS WITH A YET UNKNOWN TYPE OF CYCLIN. CAN PHOSPHORYLATE
 CC HISTONE H1.
 CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC2/CDCX SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X66357; CAA47001.1; -;
 DR PIR; S23382; S23382;
 DR PDB; 1LFN; 24-APR-02.
 DR Genew; HGNC:1772; CDK3.
 DR GK; Q00526; -;
 DR MIM; 123828; -;
 DR GO; GO:0004693; F: cyclin-dependent protein kinase activity; TAS.
 DR GO; GO:0007048; P: oncogenesis; TAS.
 DR GO; GO:0000074; P: regulation of cell cycle; TAS.
 DR InterPro; IPR000719; Prot kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
 KW Cell division; Mitosis; Phosphorylation; 3D-structure.
 KM DOMAIN 4 286 PROTEIN KINASE.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 127 127 BY SIMILARITY.
 FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 160 160 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 305 AA; 35045 MW; 1188BEC096AE462A CRC64;
 Query Match 35.0%; Score 581; DB 1; Length 305;
 Best Local Similarity 39.8%; Pred. No. 1.2e-36;
 Matches 123; Conservative 66; Mismatches 96; Indels 24; Gaps 8;
 QY 1 MEKYEKLAKTGGSGVGVFKCRNKTSQGVAVKFE---VESEDDPVVKIALREIMLKQ 57
 DB 1 MDNFQVKEKIGETGVGVTKAKNRRTGQVAVKLRLEMEGV---STAIRSEISLKE 57
 QY 58 LKHPN:VNLIIEVFRKRKMKMLVFVEYCDHTLLNLELRNPNFNG-VADGVKSVLWQTLQALNF 116
 DB 58 LKHPN:VRLD:VHNERKLYLVFEF:SQDLKXKYSTGSE:PLHLIKSYLFLQLQGVSF 117
 QY 117 CHINCHIRDRKIPENLITKGLIKICDPGPAQL-IFGDAYTDVATRWTRABELVGD 175
 DB 118 CHSHRVHRLDLPQNLINELGAIKLADFLARAFGVPLRTYTHVWTLWTRAPEILLGS 177
 QY 176 TQYSSVDVIAICVFAELLTGPLWPKGSVDVQLYLIRTLGKLIPLRHQISIFKSNGFH 235
 DB 178 KFTTAVD:VSGICIPAEVWTRKALFPDGEISDQLFRFRMLGT-----PSEDTWP 228
 QY 236 GIS-IFE-----PE-DMETLEKFSVDHPVALNFWKGLCKMPPDRITCSQLLESSEYFDS 288

Db 61 PNIVRLHDVIRKRLTLVFEYLDQDLKKYLDECGEISKPTIKSFYQLLKGVAFCCHD 120

QY 121 NCIRPDIKIBENILITKQIKICDFGAQIL-IPCDATVGVATRWYRAPELLGDTQYG 179

Db 121 RVLHDLKPNQLILNRKGLKLADEGLARAFIPVITVSHVFWTLWYRAPDLVMSGRYS 180

QY 180 SSVDIWAIGCVFAELLTQQLWPGKSDVDQYLIIRTLG-----KLIPRHQSIF- 228

Db 181 TIDWISALC:FAEMASGRPLPFGSGTSDQLFRIFKILGTPNEESWPSI:ELPEYKTDFF 240

QY 229 -----KSGFFHGISIPEDMETLEEFSDVHPVAFNFMKGCWMPDRLTCSQJES 283

Db 241 VHPAQQLSSIVEG-----LDEK-----GLNLSKYLQYDPNCRITAAALKE 282

QY 284 SYFDSFQ 290

Db 283 PYFDGLE 289

RESULT 5

CDK2_XENLA STANDARD; PRT; 297 AA.

AC P23437:

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cell division protein kinase 2 (EC 2.7.1.-) (CDK2 homolog EGI protein kinase).

GN EGI.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

CC NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93345457; PubMed=8393783;

RA Poon R.Y.C., Yamashita K., Adamczewski J.P., Hunt T., Shuttleworth C.;

RA Camonis J., Macneil S., Philippe M.;

RT "Cloning by differential screening of a Xenopus cDNA coding for a protein highly homologous to cdc2";

RL Proc. Natl. Acad. Sci. U.S.A. 88:1039-1043 (1991).

RP PHOSPHORYLATION OF THR-160.

RX MEDLINE=93345457; PubMed=8393783;

RA Poon R.Y.C., Yamashita K., Adamczewski J.P., Hunt T., Shuttleworth C.;

RT "The cdc2-related protein p40MO15 is the catalytic subunit of a protein kinase that can activate p33cdc2 and p34cdc2";

RL EMBO J. 12:3123-3132 (1993).

CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.

CC -!- INTERACTS WITH CYCLINS A, B, OR E.

CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY SIMILARITY).

CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED IN UNFERTILIZED EGG, BUT NO LONGER MADE IN THE EARLY EMBRYO.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. CDK2/CBKX SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X14227; CAA32443.1; -

DR EPR: A37871; A37871.

DR HSSP: P24941; 13UH.

DR InterPro: IPR000719; Prot_kinase.

DR InterPro: IPR002290; Ser_thr_kinase.

DR Pfam: PF00069; pkinase; 1.

DR ProDom: PD000001; Prot_kinase; 1.

DR SMART: SM00220; S_TKc; 1.

DR PROSITE: PSC0107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PSC0108; PROTEIN_KINASE_ST; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle; Cell division; Mitosis; Phosphorylation.

KW DOMAIN 4 286 PROTEIN KINASE.

FT NP_BIND 10 18 ATP (BY SIMILARITY).

FT BINDING 33 33 ATP (BY SIMILARITY).

FT ACT_SITE 127 127 BY SIMILARITY.

FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).

FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).

FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK).

SQ SEQUENCE 297 AA; 33969 MW; EC30204FCB8D198C CRC64;

Query Match: 34.4%; Score 571.5; DB 1; Length 297;

Best Local Similarity 38.1%; Pred. No. 5.8e-36;

Matches 117; Conservative 65; Mismatches 84; Indels 41; Gaps 6;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGOVAVKXFESEDDPVVKKIALREIRMLKOLKH 60

Db 2 MENFQVEKIGEGTYGVVYKARNRETGEIVALKIRLDTETEGVPSTAIRISLLKELNH 60

QY 61 PNLVNLIEVFRKRKHVLVFEYCDHTLLNELER-NPNGVADGVTKSVLWQTQLALNFCHI 119

Db 61 PNIVKLLDVHTENKLYLVFEFLNQDLKKFMDRNSIGSLVSKSYLFOLLOGLAFCHS 120

QY 120 HNCIHRDIKPNILITKOGI:KICDFGAQIL-IPGDATVGVATRWYRAPELLGDTQY 178

Db 121 HRVLHDLKPNQLILNRSOGAKLADGLARAFGVPTFTHEVTLWYRAPELLGCKFY 180

QY 179 GSSVD:WAIGCVFAELLTQQLWPGKSDVDQYLIIRTLG-----KLIPRHQSIF 228

Db 181 STAVDIWSLGCIFAEIMTRRALFGDSEIDQLFRIFRTLGTDPDEVSPGVSTTMDPKSTF 240

QY 229 KSGFFHGISIPEDMETLEEFSDVHPVAFNFMKGCWMPDRLTCSQJESYFDS 288

Db 241 P-----KWIQRDFSKVVP-----PLDEGRDLLAQMLOQ---YDS 271

QY 289 FOEAQIK 295

Db 272 NKRIKSAK 278

RESULT 6

CDK2_CARAU STANDARD; PRT; 302 AA.

ID CDC2_CARAU

AC P51958:

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).

DE CDC2.

OS Carassius auratus (Goldfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Carassius.

CC NCBI_TaxID=7957;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Oocyte;

RC Kajuru H., Yamashita M., Katsu Y., Nagahama Y.;

RA "Isolation and characterization of goldfish cdc2, a catalytic component of maturation-promoting factor.";

RL Dev. Growth Differ. 35:647-654 (1993).

CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC COMPONENT OF MPF.

```

CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN
CC MATURE OOCYTES.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D17758; BAA04605.1; --
CC FIRM: I50474; I50474.
CC HSP: Q00534; I518.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TKc_1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
CC Transferase: Serine/threonine-protein kinase; ATP-binding.
CC Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
CC FT DOMAIN 4 287 PROTEIN KINASE
CC FT BINDING 10 18 ATP (BY SIMILARITY).
CC FT ACT_SITE 33 33 ATP (BY SIMILARITY).
CC FT MOD_RES 128 128 BY SIMILARITY.
CC FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC SEQUENCE 302 AA; 34499 MW; 58DB812E19311F5 CRC64;
CC
CC Query Match 34.3%; Score 570.5; DB 1; Length 302;
CC Best Local Similarity 38.6%; Pred. No. 7e-36;
CC Matches 120; Conservative 68; Mismatches 50; Indels 33; Gaps 7;
CC
CC QY 1 MEKYLKATGEGSYGVVFKCRNKTSGOVAVKVFVESEDDPVVKKIALREIRMLKGLX 59
CC Db 1 MDYLKIEKGEGSYGVVFKGNKTTGOVAVKVKIRLESEEG-VFSYAVREISLJKELO 59
CC
CC QY 60 HPVLNVLIEVFRKRKMHVFEYCDHTLNLNELRNPG--VADGVKSVLWVTLQALNFC 117
CC Db 60 HPNVRLDLVLMQESKYLIVFEFSLMDLKKYLDSPSGQFNDPMVKSIVYLCILEGILFC 119
CC
CC QY 118 HIHNCVHRDIKPNILITKGIKICDFGFAQIL-IPGDYTDYVATWYRAPELVGDT 176
CC Db 120 HCRVLRDLKPNKLLDNKGVTKLADFGLAAGFVGVVYTVREVTLMYRAPELVLLGAS 179
CC
CC QY 177 QYSSVDIWAIGCVFAELITGQPLMPKSDVQVLYLIIRTLG-----KLIPRHQS 226
CC Db 180 RYSTPVDVMSIGTFIAELATKPLFHGDSIDQLFRIFRLTGNNVWPDVESIPDYKN 239
CC
CC QY 227 IP---KSNOFFHGISIPEDMETLEEKSDVHVFALNPMKGLCKNPPDRITCSQLLES 283
CC Db 240 TFPKWSG-----NLASTVKNLDKNGIDLLTKVLYIDPPKRSARQAVTH 284
CC
CC QY 284 SYFDSFQEAQI 294
CC Db 285 PYFDDLDKSTL 295
CC
CC RESULT 7
CC ID_CC21_ORYSA STANDARD; PRT; 294 AA.
CC AC P29618;
CC DT 01-APR-1993 (Rel. 25, Created)

```

```

DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CD2 division control protein 2 homolog 1 (BC 2.7.1.-).
GN CDC2-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, Nipponbare;
RA MEDLINE=92293101; PubMed=1376401;
RA Hashimoto C., Hirabayashi T., Hayano Y., Hata S., Chashi Y.,
RA Suzuki I., Utsugi T., Toh-E A., Kikuchi Y.;
RA "Isolation and characterization of cDNA clones encoding cdc2
RT homologues from Oryza sativa: a functional homologue and cognate
RT variants.";
RT Mol. Gen. Genet. 233:10-16(1992).
RL
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X60374; CAA42922.1; --
CC FIRM: S22440; S22440.
CC HSP: P24941; IHCL.
CC Gramenc; P29618; --
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TKc_1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
CC Transferase: Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; Phosphorylation.
CC FT DOMAIN 4 287 PROTEIN KINASE
CC FT NP_BIND 10 18 ATP (BY SIMILARITY).
CC FT BINDING 33 33 ATP (BY SIMILARITY).
CC FT ACT_SITE 127 127 BY SIMILARITY.
CC FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC SEQUENCE 294 AA; 34071 MW; 51322D93AEF4C131 CRC64;
CC
CC Query Match 34.2%; Score 568.5; DB 1; Length 294;
CC Best Local Similarity 37.6%; Pred. No. 9.6e-36;
CC Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;
CC
CC QY 1 MEKYLKATGEGSYGVVFKCRNKTSGOVAVKVFVESEDDPVVKKIALREIRMLKGLX 60
CC Db 1 MEQYKEKEKIGEGTYGVVFKRDKVTNRTIALKRLSEDEGVFSTAIRLSLKEHH 60
CC
CC QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLNL-----ELERNPNCVADGVKSVLWVTLQAL 114
CC Db 61 GNIVRLHDVHSEKRYLVFEYLDLKKFMDSCPEFAKNT-----LIKSLYQLRGV 115
CC
CC QY 115 NFCHHNCVHRDIKPNILITKQ-GIIKICDFGFAQIL-IPGDYTDYVATWYRAPELV 172

```



```
DB 116 AYCHSHRVLRLDLPONLLIDRTNALKLADFLARAFG:PVRTFTHEVWTLWYRAPEIL 175
QY 173 VGDYTGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGLKLIQRHQIFKSNQ 232
DB 176 LGSROYSPVDMWSVGCIFAENWVKPLFPDSE:DELFKIFRVLG--TFNEQS----- 227
QY 233 PFHGI-STP-----EPEDMETLEKSFSDVHPVAJNFKMGCLQNM?PDRJTCQJLE 282
DB 228 -WPGVSSLPDYKSAFFKCAQDLATI---VPTLDPAGLLSKVLREVPNKRIARQALE 283
QY 283 SSYFDSFOEAQ 233
DB 284 HEYFKDLNVQ 234

RESULT 8
CC2_ORYJA STANDARD; PRT; 303 AA.
AC Q9DGA2; Q9DGA2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein
DE kinase) (Cyclin-dependent kinase 1) (CDK1).
GN CDC2.
OS Oryzias javanicus (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID:123683;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yamashita M., Mita K.;
RT "cDNA cloning of Cdc2 and cyclin B in medaka species.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBA databases.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL
CC CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND
CC MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE
CC REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC
CC COMPONENT OF MPF (By similarity).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN
CC NATURE OCCTES (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/arnoc/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB050461; BAB17219.1; -
CC EMBL; AB050462; BAB17220.1; -
CC HSSP; Q00534; 1B18.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; pkinase.1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SMC0220; S_IK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
CC DOMAIN 4 287 PROTEIN KINASE.
```

```
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT VARIANT 49 49 V -> I.
FT VARIANT 245 245 M -> K.
FT VARIANT 269 269 I -> T.
FT VARIANT 276 276 I -> V.
FT VARIANT 299 299 C -> S.
SQ SEQUENCE 303 AA; 3469; MW; 43988B012BE28D9C CRC64;

Query Vabch. 34.1%; Score 567.5; DB 1; Length 303;
Best Local Similarity 37.9%; Pred. No. 3.2e-35;
Matches 117; Conservative 72; Mismatches 31; Indels 29; Gaps 7;

QY 1 MEKYEKLAKTGEGSYGVYFKENKTSQGVAVKXP-VESEDDPVVKKIALREIRMLKOLK 59
DB 2 MEDYVKIEKIGEGTYGVYKHKSTQGVAMKKIRLESEEG-VPTAVREVSLQELK 59
QY 60 HENLVNLEIEVFRKRRKHVFEYCDHTLLNELERNPG--VADGVIKSVLWCTLOALNFC 117
DB 60 HENVRLLDVLMQESRLYLIFEFLSMDLKKYLDSTPSQVYMDPMLVKSYLQILEGIYFC 119
QY 118 H-HNCHIRDKPENILITKQG-IKIDFGFAQ-L-IPGAYTVYVATWYRAPELLVGST 176
DB 120 HRRVRLHRLDKPQLNLLDKGVKIKLADFLARAFGVPVRVTVTHEVTLWYRAPEVLGSP 179
QY 177 QYGVSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLG-----KLTPRHQS 226
DB 190 RYSTPVVWSTGTIFAELATKAPLPHGSEIDQLPRIPTLGTNNVDPVDPVES-PDYKN 239
QY 227 IFKSNQPFHGISIPPEQNE-TLEKFSVDVHPVALNFKYKGLQNMPPDRRLTCSQ-LESSY 285
DB 240 TP-----PKWEGSLSSWVKNLDKGLLAKMLIYNPPKRIAREANTHPY 286
QY 286 FDSFQEAQI 294
DB 287 FDDLKSTL 295

RESULT 9
CDK2_HUMAN STANDARD; PRT; 298 AA.
AC P24941;
DT 01-MAR-1992 (Rel. 21, Created);
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.-) (p33 protein kinase).
GN CDK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91330891; PubMed=1714386;
RA Elledge S.J., Spottswood M.R.;
RT "A new human p34 protein kinase, CDK2, identified by complementation
RT of a cdc28 mutation in Saccharomyces cerevisiae, is a homolog of
RT Xenopus Egl.";
RL EMBO J. 10:2653-2659(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91367262; PubMed=1653904;
RA Tsai L.-H., Harlow E., Meyerson M.;
RT "Isolation of the human cdk2 gene that encodes the cyclin A- and
RT adenovirus E1A-associated p33 kinase.";
RN Nature 353:174-177(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020980; PubMed=1717994;
```

RA Ninomiya-Tsuji J., Nomoto S., Yasuda H., Reed S.I., Matsumoto K.;
RT "Cloning of a human cDNA encoding a CDC2-related kinase by
RL complementation of a budding yeast cdc28 mutation";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9006-9010(1991).
RN [4]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,
RA Krausner R.D., Collins F.S., Wagner L., Shenren C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.J., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Valek J.A., Gunatirne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.C., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Szallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP PHOSPHORYLATION SITES.
RX MEDLINE=93010995; PubMed=1396589;
RA Gu Y., Rosenblatt J., O'Morgan D.O.;
RT "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160
RL and Tyr15";
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=93288132; PubMed=8510751;
RA de Bondt H.L., Rosenblatt J., Jancarik J., Jones H.D.,
RA Morgan D.O., Kim S.-H.;
RT "Crystal structure of cyclin-dependent kinase 2";
RL Nature 363:595-602(1993).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
RX MEDLINE=95356811; PubMed=7630397;
RA Jeffrey P.D., Russo A.A., Polyak K., Gibbs E., Hurwitz J.,
RA Massague J., Pavietich N.P.;
RT "Mechanism of CDK activation revealed by the structure of a
RL cyclinA-CDK2 complex";
RL Nature 376:313-320(1995).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
RX MEDLINE=96181416; PubMed=8610110;
RA de Azevedo W.F. Jr., Mulder-Dieckmann H.-J., Schulze-Gahmen U.,
RA Worland P.J., Sausville E., Kim S.-H.;
RT "Structural basis for specificity and potency of a flavonoid
RL inhibitor of human CDK2, a cell cycle kinase";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2735-2740(1996).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH KIP1.
RX MEDLINE=96300318; PubMed=8684460;
RA Russo A.A., Jeffrey P.D., Patten A.K., Massague J., Pavietich N.P.;
RT "Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor
RL bound to the cyclin A-Cdk2 complex";
RL Nature 382:325-331(1996).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
RX MEDLINE=96313126; PubMed=8756328;
RA Russo A.A., Jeffrey P.D., Pavietich N.P.;
RT "Structural basis of cyclin-dependent kinase activation by
RL phosphorylation";
RL Nat. Struct. Biol. 3:696-700(1996).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97075215; PubMed=8917641;
RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;
RT "High-resolution crystal structures of human cyclin-dependent kinase
RL 2 with and without ATP: bound waters and natural ligand as guides for
RL inhibitor design";
RL J. Med. Chem. 39:4540-4546(1996).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=97475219; PubMed=9334743;
RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,
RA Endicott J.A.;
RT "Protein kinase inhibition by staurosporine revealed in details of
RL the molecular interaction with CDK2";
RL Nat. Struct. Biol. 4:796-801(1997).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CKS1.
RX MEDLINE=96182647; PubMed=8601310;
RA Bourne Y., Watson M.H., Hickey M.J., Holmes W., Rocque W., Reed S.I.,
RA Tainer J.A.;
RT "Crystal structure and mutational analysis of the human CDK2 kinase
RL complex with cell cycle-regulatory protein CksH1";
RL Cell 84:863-874(1996).
RN [15]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=98342389; PubMed=9677190;
RA Gray N.S., Wodicka L., Thummlen A.-M.W.H., Norman T.C., Kwon S.,
RA Espinoza F.H., Morgan D.O., Barnes G., Leclerc S., Mettler L.,
RA Kim S.H., Lockhart D.J., Schultz P.G.;
RT "Exploiting chemical libraries, structure, and genomics in the search
RL for kinase inhibitors";
RL Science 281:533-538(1998).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLIN A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; X61622; CAA43807.1; -;
DR EMBL; X62071; CAA43985.1; -;
DR EMBL; M68520; AAA35667.1; -;
DR EMBL; AF512553; AAM34794.1; -;
DR EMBL; EC003065; AA03065.1; -;
DR PIR; A41227; A41227;
DR PDB; 1F2N; 27-JAN-97.
DR PDB; 1HCK; 07-DEC-96.
DR PDB; 1HC2; 07-DEC-96.
DR PDB; 1A01; 12-NOV-97.
DR PDB; 1JST; 11-JAN-97.
DR PDB; 1J50; 29-JUL-97.
DR PDB; 1B0H; 09-SEP-98.
DR PDB; 1B38; 23-DEC-98.
DR PDB; 1B39; 23-DEC-98.
DR PDB; 1CKP; 13-JAN-99.
DR PDB; 1D18; 28-JAN-03.
DR PDB; 1DM2; 31-MAY-00.
DR PDB; 1E1V; 10-MAY-01.

```

DR PDB: 1E1X; 10-MAY-01.
DR PDB: 1F5Q; 27-DEC-00.
DR PDB: 1FQ1; 09-MAY-01.
DR PDB: 1FV1; 17-JAN-01.
DR PDB: 1FV2; 17-JAN-01.
DR PDB: 1G5S; 14-NOV-01.
DR PDB: 1G1H; 06-NOV-02.
DR PDB: 1G1I; 06-NOV-02.
DR PDB: 1G1J; 06-FEB-02.
DR PDB: 1G1K; 06-JUN-02.
DR PDB: 1H1P; 19-SEP-02.
DR PDB: 1H1Q; 19-SEP-02.
DR PDB: 1H1R; 19-SEP-02.
DR PDB: 1H1S; 19-SEP-02.
DR PDB: 1H24; 01-FEB-03.
DR PDB: 1H25; 01-FEB-03.
DR PDB: 1H26; 01-FEB-03.
DR PDB: 1H28; 01-FEB-03.
DR PDB: 1J5V; 29-AUG-01.
DR PDB: 1JVP; 21-DEC-01.
DR PDB: 1KE5; 14-MAY-02.
DR PDB: 1KE6; 14-MAY-02.
DR PDB: 1KE7; 14-MAY-02.
DR PDB: 1KE8; 14-MAY-02.
DR PDB: 1KE9; 14-MAY-02.

Query Match 34.1%; Score 567; DB 1; Length 298;
Best Local Similarity 40.3%; Pred. No. 1.3e-35;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTCEGSGVGVKCRNKTSGOVAVKVFESDDPVVVKALREIMLKOLKH 60
DB 1 MENFQKVEKIGETGVGVVYKARNKLTGEVWALKIRLQTEGVPSTAIRSILKEMNH 60
QY 61 PNLVNLIEVFRKRKRKHLEVFYCDHTLNLN-NGVADGV-KSVLKYTLQALNFCHE 1-9
DB 61 PNLVNLIEVFRKRKRKHLEVFYCDHTLNLN-NGVADGV-KSVLKYTLQALNFCHE 120
QY 120 NCHHRCIKPENILITKQ-GIHKICDFGPAQIL-IPGDAYTDYVATRWYRAPELL 178
DB 121 HRVLRHDLKPKNLLTGAIKLADFLGRLAFGIPVTRTTHVWTLWYRAPEILLGCKY 180
QY 179 GSSVDIMWAGCVFAELLTGOPLWPKGSVDVQLYLIIRTLGLKLIPIHQSIKFSNG 237
DB 181 STAVDIWSLGCIFAEVYTRALFPDSEIDQLFRIFRLGT-----PDEVWVEGVT 231
QY 238 SIPE--PEMETLEKFSVDHPV-----ALNFMKGLKKNPDRLTCSQLLESSYF 286
DB 232 SMPDYKSPFKWARQDFSKWVPLDEGDSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 10
CDC2_MAIZE STANDARD; PRT; 294 AA.
AC P23111;
CT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
ET 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34cdc2).
GN CDC2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91195354; PubMed=2014238;
RA Colasanti J., Tyers M., Sundaresan V.;
RT "Isolation and characterization of cDNA clones encoding a functional
RT p34cdc2 homologue from Zea mays."
RL Proc. Natl. Acad. Sci. U.S.A. 88:3377-3381(1991).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC

```

```

CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M60526; AAA33479.1; -.
CC HSSP: P24941; 1HCL.
CC MaizeDB: 60686; -.
CC
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR002290; Ser_thr_kinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
CC Trasnferase: Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; phosphorylation.
CC FT DOMAIN 4 287
CC FT NP_BIND 10 18 ATP (BY SIMILARITY).
CC FT BINDING 33 33 ATP (BY SIMILARITY).
CC FT ACT_SITE 127 127 BY SIMILARITY.
CC FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC SQ SEQUENCE 294 AA; 33934 MW; 5063ECFC2D5FFDD CRC64;

Query Match 34.1%; Score 566.5; DB 1; Length 294;
Best Local Similarity 37.6%; Pred. No. 1.4e-35;
Matches 117; Conservative 66; Mismatches 93; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTCEGSGVGVKCRNKTSGOVAVKVFESDDPVVVKALREIMLKOLKH 60
DB 1 MEQYEKVEKIGETGVGVVYKALDKATNIIALKIRLEQDEGVSTAIRSILKEMNH 60
QY 61 PNLVNLIEVFRKRKRKHLEVFYCDHTLNLN-----EERNPNGVADGVKSVLKYTLQAL 114
DB 61 GNIVLRHDLVHSEKRIYLVFVFLDLDLKKPMDSCPEFAKNPT----LIKSYLYQLHGV 115
QY 115 NFCHHRCIKPENILITKQ-GIHKICDFGPAQIL-IPGDAYTDYVATRWYRAPELL 172
DB 116 AYCHSHRVLRHDLKPKNLLIDRTNALKLADFLGRLAFGIPVTRTTHVWTLWYRAPEIL 175
QY 173 VGDITQYSSVDIMWAGCVFAELLTGOPLWPKGSVDVQLYLIIRTLGLKLIPIHQSIKFSNG 232
DB 176 LGARQYSTPVDVWSVGCIFAEVWQKPLFPDSEIDELFKIFRLG--TPNEQS----- 227
QY 233 PPHGIS-IP-----EPDMETLEKFSVDHPVVALNFMKGLKKNPDRLTCSQLLE 282
DB 228 -WPGVSCLPDFKTAFFRWQAQDLATV---VPLNDPAGLDLILSKMLRVEPSKRITARGALE 283
QY 283 SSYPDSFOEAQ 293
DB 284 HEYFKDLEVVQ 294

RESULT 11
CDC2_CRIGR
ID -CDC2_CRIGR STANDARD; PRT; 298 AA.
AC O55076;
ET 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

```


RA Mayer K.F.X., Kaul S., Town C.D., Koc H.J., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Miltscher J., Sellers P., Gili J.E., Feldhlyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Karko T., Nakamura Y., Sato S., Kato T., Asanizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RL thaliana";
CC Nature 408:820-822(2000).
CC
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC -!- CAUTION: Ref.5 sequence differs from that shown: due to erroneous
CC gene model prediction.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; S45387; AB223643.1; -;
CC EMBL; Y59198; AA328331.1; -;
CC EMBL; D19850; BA016231.1; -;
CC EMBL; X57839; CA44971.1; -;
CC EMBL; AL132963; CAB87903.1; ALT_SEQ.
CC PIR; S23095; S23095.
CC HSP2; P24941; 1HCL.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SMC0220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
CC Cell division; Mitosis; Phosphorylation.
CC KW DOMAIN 4 287 PROTEIN KINASE.
CC FT NP_BIND 10 18 ATP (BY SIMILARITY).
CC FT BINDING 33 33 ATP (BY SIMILARITY).
CC FT ACT_SITE 127 127 BY SIMILARITY.
CC FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC SQ SEQUENCE 294 AA; 34030 MW; B5FAE55PA9EC366E CRC64;

Query Match 33.8%; Score 561.5; DB 1; Length 294;
Best Local Similarity 37.1%; Pred. No. 3.2e-35;
Matches 111; Conservative 70; Mismatches 93; Indels 25; Gaps 6;

Qy 1 MEKYEKLAKTGEYGVVGVCKRNKTSQGVAVKVFSESDPPVKKIALREIRMLKQLKH 60
Db 1 MDQYKVEKIGEGYGVVYKARDKVNTETALKIRLEQDEGVPSTAIRISLLKEMQH 60
Qy 61 PNVNLEIVFRKRKRLHVEYCDHTLLNELRNPNGVAD-GVKSVLWOTQALNFCH 119
Db 61 SNIVKLQDVHSEKRLVLYEYLDLKGMDSPDFSKLHMKYLYQILSGAYCHS 120
Qy 120 HNCIHRRIKPNLITKQ-GIIKICDRGFAQIL-IPGQNTVDVATRWYRPELLVGQTC 177
Db 121 HRVLHRLKPNQLIDRRTNSKLADFLGARAFGIPVRITTHEVYTLWYRPEILLGSHH 180

QY 178 YGSSVDIWAIGCAVFAELITGQPLWPGKSDVDQLYLIIRTLG-----KLIPRHQSI 227
Db 161 YSTPVDIWSVGCIFAEMISQKPLFGSDIDQLFKIFIMGTPVEDTWRGVTSLPDYKSA 240
QY 228 PKNSGFFHGISIPEDMETLEERFSDVHPVALNPMKGCCKNPDRLTCSQLLESSYF 286
Db 241 FPK-----WKPTOLETF---VPLNDPDGVLLSKMLMLMDPTKRINARAALAEHYF 287

RESULT 13

CDK2_ORYLA
ID CDC2_ORYLA STANDARD: PRT; 303 AA.
AC Q9DGS3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein
DE kinase) (cyclin-dependent kinase 1) (CDK1).
DE CDC2.
GN Oryzias latipes (Medaka fish) (Japanese ricefish).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yokota T., Yoshida N., Matsui H., Takahashi T., Yamashita M.;
RT "cDNA cloning of medaka (Oryzias latipes) Cdc2".
CC Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL
CC CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND
CC MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE
CC REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC
CC COMPONENT OF MPF (By similarity).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN
CC MATURE OOCYTES (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB040436; BAB13720.1; -;
CC HSP; Q00534; 1B18.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SMC0220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
CC KW DOMAIN 4 287 PROTEIN KINASE.
CC FT NP_BIND 10 18 ATP (BY SIMILARITY).
CC FT BINDING 33 33 ATP (BY SIMILARITY).
CC FT ACT_SITE 128 128 BY SIMILARITY.
CC FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).

```
SQ SEQUENCE 303 AA; 34688 MW; 43988D072B8E4889C CRC64;
Query Match 33.8%; Score 561.5; DB 1; Length 303;
Best Local Similarity 37.5%; Pred. No. 3.6e-35;
Matches 116; Conservative 72; Mismatches 92; Indels 29; Gaps 7;
QY 1 MEKYELAKTGGSGYGVFKCRNKTSQGVAVKVF-VESEDDPVVKKIALREIMLKQK 59
DB 1 MEDYVKEIKEGTGVVYKGRHKSTQGVAMKKIRLESEEG-VPSTAVREVSILQELK 59
QY 60 HNLNVLTEVPRKRMHVLFEYCDHTLLNELRNPNNG--VADGVKSVLQTLQANFC 117
DB 60 HNVVRLDLVLMQERLYLIFELSLMDLKKYJDSIPSGYNDPMLVKSYLEQLEGYFC 119
QY 118 HTHNCIHRDIPENILITKQGIKICDFGAQL-IPGCAVTEYVATRWYRAPELJXGDT 176
DB 120 HRRVLRDLKQNLIDNKGVIKLADEGLARAFGVPRVTVTHEVTLWYRAPELVLLGSP 179
QY 177 QYSSVCIWAICVFAELLTGPFGKSDVDCLYLITLGL-----KLIPRQHS 226
DB 180 RYSTVDVMSTGTFAELATKPKFPHGSEIDQLFRIFRTLTPNNQVMPDVEGLPDYKN 239
QY 227 IFKSGFFHGHSIPPEDEVE-TLEEFSDVHVPVALKEMKGLCKMPDRLTCSOLLESSY 285
DB 240 TF-----PKWKEGSLSSWVNLKNGLDLLAKMLIYNPKRISAREAYTHPY 286
QY 286 FDSFOEAQI 294
DB 287 FDDLKSTL 295
RESULT 14
CDK2_RAT
ID CDK2_RAT STANDARD; PRT; 298 AA.
AC Q63699; O09136;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.-).
GN CDK2.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=95166553; PubMed=8673024;
RA Kotani S., Endo T., Kitagawa M., Higashi H., Onaya T.;
RT "A variant form of cyclin-dependent kinase 2 (cdk2) in a malignantly
transformed rat thyroid (FRTL-Tc) cell line."
RL Oncogene 10:663-669(1995);.
RN [2]
RP SEQUENCE OF 19-124 FROM N.A.
RX MEDLINE=96113578; PubMed=8673024;
RA Hosokawa Y., Yang M., Kaneko S., Tanaka M., Nakashima K.;
RT "Synergistic gene expressions of cyclin E, cdk2, cdk5 and E2F-1
during the prolactin-induced G1/S transition in rat B62 pre-T
lymphoma cells."
RL Biochem. Mol. Biol. Int. 37:393-399(1995).
CC Biochem. Mol. Biol. Int. 37:393-399(1995).
CC !- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
DURING S PHASE AND G2.
CC !- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
SIMILARITY).
CC !- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=CDK2 alpha;
CC isoId=Q63699-1; Sequence=Displayed;
CC isoId=Q63699-2; Sequence=Not described;
CC !- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC
```

```
CC CDC2/CDKX SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; D28753; BAA05947.1; -
DR EMBL; D63162; BAA09638.1; -
DR HSSP; P24941; IAQ1.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR Transfaser; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Cell division; Mitosis; Phosphorylation; Alternative splicing.
FT DOMAIN 4 286 PROTEIN KINASE
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT CONFLICT 79 79 V -> C (IN REF. 2).
FT CONFLICT 99 99 L -> I (IN REF. 2).
FT CONFLICT 124 124 L -> C (IN REF. 2).
SQ SEQUENCE 298 AA; 33887 MW; C8CB3ADCE9B97F88 CRC64;
Query Match 33.8%; Score 561; DB 1; Length 298;
Best Local Similarity 39.7%; Pred. No. 3.6e-35;
Matches 117; Conservative 61; Mismatches 99; Indels 18; Gaps 6;
QY 1 MEKYELAKTGGSGYGVFKCRNKTSQGVAVKVFVESEDDPVVKKIALREIMLKQK 60
DB 1 MENFQVKEIKEGTGVVYKAKNKLTEGVVALLKRLDTETEGVPTAIRISLLKELNH 60
QY 61 PNLVNLIEVPRKRMHVLFEYCDHTLLNELRNPN-NGVADGVKSVLQTLQANFC 119
DB 61 PNIVKLDVHTENKLDYVFEFLHQDLKFKFMDASALTGLPLIKSYLFQQLQGLAFCHS 120
QY 120 HNCIHRDIPENILITKQGIKICDFGAQL-IPGCAVTEYVATRWYRAPELJXGDTQY 178
DB 121 HRYVLRDLKQNLINAEGLKADFLARAFGVPRVTVTHEVTLWYRAPELJXGDTQY 180
QY 179 GSSVDIWAICVFAELLTGPFGKSDVDCLYLITLGLKLIPIRQHSIFKSGFFHGI- 237
DB 181 STAVD:WSLGCIPAEVMTTRALFPQDSE:DQLFRIPRTLTG-----PDEVVMPGVT 231
QY 238 SIPE--PEDMETLEEFSDVHPV----ALNFMKGLCKMPDRLTCSOLLESSYF 286
DB 232 SMPDYKPSFKMARQDFSKVPPPLEDGRSLSSQMLHYDPNKRIRISAKAALAHPPF 286
RESULT 15
CDK2_ORYCU
ID CDK2_ORYCU STANDARD; PRT; 303 AA.
AC Q9DGA5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein
kinase) (Cyclin-dependent kinase 1) (CDK1).
GN CDK2.
OS Oryzias latipes (Hymenoptera).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
```

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=104658;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yamashita M., Mita K.;
RT "cDNA cloning of Cdc2 and cyclin B in medaka species."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL
CC CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND
CC MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE
CC REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC
CC COMPONENT OF MPF (By similarity).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME. WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN
CC MATURE OOCYTES (by similarity).
CC -!- SUBCELLULAR LOCATION: Nucleus (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/acnucnue/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB050458; BAB17216.1; -;
DR HSP; Q00534; 18-8.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser/Thr_pkinase.
DR Pfam: PF00669; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART; SMC0220; S_TKc; 1.
DR PROSITE; PSC0107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSC0111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PSC0108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Nucleic acid; Phosphorylation.
FT DOMAIN 4 287 PROTEIN KINASE
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 303 AA; 34605 MW; 65ECD9A98924888F CRC64;
Query Match 33.7%; Score 559.5; DB 1; Length 303;
Best Local Similarity 37.5%; Pred. No. 4.7e-35;
Matches 116; Conservative 71; Mismatches 93; Indels 29; Gaps 7;
QY 1 MEKYEKATKGGSGVWFKCNKTSQGVAVKKF-VSEDDPVVKKIALREIRMLKOLK 59
DB 1 MEDYVKIEKIGGTGVVYKGRHKSTQGVANWKIRLESEEG-VPSTAVREVSLLOELK 59
QY 6C HENLVNLTIEVFRKRKCHLVFECYDHTLLNE-LERNPNG--VADGVKSVLWQTLQALNFC 117
DB 60 HENVVRLLOLVNCSRLYLIFELSLMDLKKYLDSPSCQYMDPMVLVKSPLYQLLEGIVFC 119
QY 118 HINCHTRDRIKPNILITKQIKICDFGAQIL-IPGAYTDYVATRWYRAPPELLVQCT 176
DB 120 HRRRVLRDLKPKQNLIDNKGVIKLADFLGSLFAFGVPRVVTHEVVT-LWYRAPEVLLGSP 179
QY 177 QYGVSSVDIWAIGCVFAELITGQPLWPKGSDVDQVLYLIRTLG-----KLIPRHQS 226
DB 180 RYSTPVDVWSTGTIFALATKKPLFHGDSIEDQLFRITLGTGPNVWPDVSESJPDYKS 239
QY 227 IF-KSNGFFHGSIPEPEDMETLEEKFSQVHFVA-LFMKGCIMKMPDSDRLTCSQ-LRSSY 285

Db 240 TFPKWKG-----CSLSMWKQLDKGLDLAKMLIYNPKKRISAREAMTHPY 286
QY 286 FDSFQEAQI 294
DB 287 FDDLKSTL 295
Search completed: November 13, 2003, 14:19:20
Job time : 10.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OV protein - protein search, using sw mode:

Run on: November 13, 2003, 14:16:06 ; Search time 29.9773 Seconds
(without alignments)
2711.600 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1662

Sequence: 1 MEKYEKLAKTGGSGYGVFK.....RKARNEGRRRQQVLPPLKS 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREXBL_23:

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_mhc:
- 8: sp_organelle:
- 9: sp_phase:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:
- 15: sp_virus:
- 16: sp_bacteriap:
- 17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1206	72.6	352 11	Q8CEQ0 mus musculus
2	1027.5	61.8	353 5	Q9U2H1 canorhachidi
3	1027.5	61.8	392 5	Q9VMN3 drosophila
4	967.5	59.2	566 6	Q9TK0 oryctolagus
5	941.5	56.6	493 4	Q92772 homo sapien
6	930.5	56.0	564 11	Q9Y12 mus musculus
7	930.5	56.0	568 11	Q9Y12 mus musculus
8	928	55.8	329 11	Q9Y11 mus musculus
9	784.5	47.2	455 4	Q9Y14 homo sapien
10	784.5	47.2	592 4	Q8YVW4 homo sapien
11	781	47.0	657 5	Q9BMG2 trypanosoma
12	770	46.3	353 11	Q8K134 mus musculus
13	769.5	46.3	1106 5	Q9GRT9 leishmania
14	767.5	46.2	457 11	Q8KXR2 mus musculus
15	764.5	46.0	585 11	Q8B149 mus musculus
16	763.5	45.9	457 11	Q9JM02 rattus norv

17	763.5	45.9	505 11	Q9JM01 rattus norv
18	759.5	45.7	595 11	Q8BLF2 mus musculus
19	720.5	43.4	783 11	Q8BW18 mus musculus
20	720.5	43.4	831 4	Q8IYC7 homo sapien
21	708	42.6	993 4	Q8WX05 homo sapien
22	702.5	42.3	578 11	Q8BYE0 mus musculus
23	644	38.7	1104 13	Q9W6R6 fugu rubrip
24	641.5	38.6	997 4	Q9UJ16 homo sapien
25	579.5	34.9	294 10	Q9ZRI1 tritium ae
26	564	33.9	288 5	Q9XZD6 plasmodium
27	562	33.8	288 5	Q96820 plasmodium
28	560	33.7	288 5	Q96821 plasmodium
29	555.5	33.4	300 3	Q13379 pneumocysti
30	555.5	33.4	300 3	Q13380 pneumocysti
31	555.5	33.4	300 5	Q15890 toxoplasma
32	554.5	33.4	294 10	Q82666 brassica na
33	554.5	33.4	294 10	Q40790 pinus conto
34	552.5	33.2	300 5	Q40000 toxoplasma
35	551	33.2	292 13	Q9DE44 brachydanio
36	550.5	33.1	294 10	Q43361 picea abies
37	549.5	33.1	294 5	Q76541 cryptospori
38	549.5	33.1	294 10	Q9FUR4 nicotiana t
39	549.5	33.1	294 10	Q40484 nicotiana t
40	549.5	33.1	297 10	Q8RX68 arabidopsis
41	549	33.0	293 10	Q8GT22 oryza sativ
42	549	33.0	298 5	Q27032 theileria p
43	547.5	32.9	294 10	Q8L6T8 daucus caro
44	547.5	32.9	294 10	P93556 sesbania ro
45	547.5	32.9	300 5	Q17066 asterina pe

ALIGNMENTS

RESULT 1

Q8CEQ0 PRELIMINARY; PRT; 352 AA.
AC Q8CEQ0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Cyclin-dependent kinase-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
PA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
DR Nature 420:563-573(2002).
DR EMBL; AK016781; BAC25497.1;
SQ SEQUENCE 352 AA; 41023 MW; 6E9F56C8080F35E0 CRC64;

Query Match 72.6%; Score 1206; DB 11; Length 352;
Best Local Similarity 68.9%; Pred. No. 2.2e-102;
Matches 222; Conservative 42; Mismatches 48; Indels 10; Gaps 3;

QY	1	MEKYEKLAKTGGSGYGVFKCRNKTSGQVAVAKFVESEDDPVVKIALREIRMLKQLKH	60
DB	1	MEKYEKIGKIGESYGVFKCRNRDTGQVAKRFLTEDDPVVKIALREIRMLKQLKH	60
QY	61	PNLVNLIEVFRKRKGLHVFYCDHTLNLERNPNGVADGVKSVLWQTLQALNFCCHI	120
DB	61	PNLVNLLEVFRKRKRLHVFYCDHTLVHLELDYRQGVPEPLVKNITWQTLQAVNFCCH	120
QY	121	NCIHRDKPENILITKQIHKICDPGFAQILI-PGDAYTDYVATRWYRAPPELLVGDTQYG	179
DB	121	NCIHRDKPENILITKQSAIKLCLDFGARLLTGPDDYTDYVATRWYRSPPELLVGDTQYG	180

QY	180	SSVDIWAIGCVFAELLTGOPLWPKGSDVDQLYLIRTLGKLIIRHOSIFKSNQFFHGISI	239	QY	241	EPEDMETLEEKFSVHPVVALNFMKGLQMNPDRLTCSQLLESSYFDSFOBAQIKRKARN	300
DB	181	PPVDIWAIGCVFAELLTGOPLWPKGSDVDQLYLIRTLGKLIIRHOSIFKSNQFFHGISI	240	DB	241	EPEHLEPLPSKLPNASSAQLDFLOKCFEMSPDRFRSCSELMHLGIFSNW----	296
QY	240	PEPEDMETLEEKFSVHPVVALNFMKGLQMNPDRLTCSQLLESSYFDSFOBA-----QI	294	QY	300	EG-----RNRQRQVLP	313
DB	241	PDPEMETLEKFNISYALGFLKGCJHMDPAERLTCEQLQHPYFDSIREVGELTRQH	300	DB	297	ESTPTGLTSKSPNPLP	314
QY	295	KRKAREGRNRQR-----QVLP	312	DB	297	ESTPTGLTSKSPNPLP	314
DB	301	DKPARKTLRQGRKLTGLQVLP	322				
RESULT 2							
Q9U2H1		PRELIMINARY;	PRT;	353	AA.		
ID	Q9U2H1						
AC	Q9U2H1						
DT	01-MAY-2000	(Tremblrel. 13, Created)					
DT	01-MAY-2000	(Tremblrel. 13, Last sequence update)					
DT	01-MAR-2003	(Tremblrel. 23, Last annotation update)					
DE	Y42A5A.4	protein.					
GN	Y42A5A.4						
OS	Caenorhabditis elegans.						
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;						
OC	Rhabditidae; Peleoderinae; Caenorhabditis.						
CX	NCBI_TaxID=6239;						
RN	SEQUENCE FROM N.A.						
RA	Steward C.A.						
RL	Submitted (Oct-1998) to the ENRL/GenBank/DBJ databases.						
RN	[2]						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=39069613; PubMed=9851916;						
RA	none;						
RT	"Genome sequence of the nematode C.elegans: A platform for						
RT	investigating biology";						
RL	Science 282:2012-2018(1998).						
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.						
DR	EWBL; ALG32618; CAB63367.1; "						
DR	HSSP; P24941; 1HCL.						
DR	WormPep; Y42A5A.4; CE20258.						
DR	InterPro; IPR000719; Prot_kinase.						
DR	InterPro; IPR002290; Ser_thr_pkinase.						
DR	Pfam; PF00069; pkinase; 1.						
DR	ProDom; PDC00001; Prot_kinase; 1.						
DR	SMART; SM00220; S_TKC; 1.						
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.						
DR	PROSITE; PS00111; PROTEIN KINASE DOM; 1.						
DR	PROSITE; PS00108; PROTEIN KINASE ST; 1.						
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.						
SQ	SEQUENCE 353 AA; 40665 MW; 7E32260D9849CAC9 CRC64;						
Query Match		61.8%;	Score 1027.5;	DB 5;	Length 353;		
Best Local Similarity		58.8%;	Pred. No. 5.4e-86;				
Matches 187;	Conservative	59;	Mismatches	63;	Indels	9;	Gaps
QY	1	MEKYEKLAKTGEYSYGVVFKRNKTSQVAVKVFVSEDDPVVKIALREIRMLKQKH	60	QY	1	MEKYEKLAKTGEYSYGVVFKRNKTSQVAVKVFVSEDDPVVKIALREIRMLKQKH	60
DB	1	MDKYDRLSKLGEGSYGVVYKCKNRDTCQVIAIKKFVETEDDPHKKIALREIRMLKQKH	60	DB	1	MDKYDRLSKLGEGSYGVVYKCKNRDTCQVIAIKKFVETEDDPHKKIALREIRMLKQKH	60
QY	61	PNLVNLLVEPRKKEKHLVPEYCDHTLLNELRPNPGVAGVTKSLVWQTLQALNFCHE	120	QY	61	PNLVNLLVEPRKKEKHLVPEYCDHTLLNELRPNPGVAGVTKSLVWQTLQALNFCHE	120
DB	61	QNLVGLIEVFRNKLHLVPELQRTVHLEKRNPHGVNDLIKIIYQLLEALKFCHSH	120	DB	61	QNLVGLIEVFRNKLHLVPELQRTVHLEKRNPHGVNDLIKIIYQLLEALKFCHSH	120
QY	121	NCIHRDKPENILITKGIKICDQFAQILIPGCAYTIVATRWAEAPELLVAGDTQYGS	180	QY	121	NCIHRDKPENILITKGIKICDQFAQILIPGCAYTIVATRWAEAPELLVAGDTQYGS	180
DB	121	KCIHRDKPENILITRNDQVKGIDFGFARIINTEMYTDYVATRWAEAPELLVAGDTQYGP	180	DB	121	KCIHRDKPENILITRNDQVKGIDFGFARIINTEMYTDYVATRWAEAPELLVAGDTQYGP	180
QY	181	SVDIWAIGCVFAELLTGOPLWPKGSDVDQLYLIRTLGKLIIRHOSIFKSNQFFHGISI	240	QY	181	SVDIWAIGCVFAELLTGOPLWPKGSDVDQLYLIRTLGKLIIRHOSIFKSNQFFHGISI	240
DB	181	PVDIWAIGCVFAELLTGOPLWPKGSDVDQLYLIRTLGKLIIRHOSIFKSNQFFHGISI	240	DB	181	PVDIWAIGCVFAELLTGOPLWPKGSDVDQLYLIRTLGKLIIRHOSIFKSNQFFHGISI	240

```

DR SMART; SMO0220; S TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 392 AA; 44962 MW; C9627D07784E5D CRC64;

Query Match 61.8%; Score 1027.5; DB 5; Length 392;
Best Local Similarity 53.4%; Pred. No. 6.6e-86;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

CY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
CY 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
DB 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
CY 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
DB 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
CY 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
DB 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
CY 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
DB 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
CY 180 SVDIWAIGCVFAELLTGQPLWPKGSVDQYLIIRTLGKLIIRHOSIFKSNCFEGHISI 239
DB 180 SVDIWAIGCVFAELLTGQPLWPKGSVDQYLIIRTLGKLIIRHOSIFKSNCFEGHISI 239
CY 181 KAVDWAIGCLVTEMLNGEPLFFGDSIDQYLIIMRCLGNLIPRHQELFYKNPVFAGVRL 240
DB 181 KAVDWAIGCLVTEMLNGEPLFFGDSIDQYLIIMRCLGNLIPRHQELFYKNPVFAGVRL 240
CY 240 PEPEDMETLBEKFDVHPVALNFMKGLKNPDDRLTCSQLLESSYF--DSF-----QEA 292
DB 240 PEPEDMETLBEKFDVHPVALNFMKGLKNPDDRLTCSQLLESSYF--DSF-----QEA 292
CY 241 PEIKSEPLERRYPKLVSEVVIDLAKKCLHVDPKRPFCAELLHDDFFQMDGFAERFSQEL 300
DB 241 PEIKSEPLERRYPKLVSEVVIDLAKKCLHVDPKRPFCAELLHDDFFQMDGFAERFSQEL 300
CY 293 Q-K--RKARN-----EGNRPRQO 309
DB 301 QMKYQKDARNISLSKKSQNRKKEK 324

RESULT 5
Q92772 PRELIMINARY; PRT; 493 AA.
ID Q92772
AC Q92772;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P56 KKIAMRE protein kinase.
GN KKIAMRE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97152547; PubMed=9000130;
RA Taglietti C.A., Wyak M., Davis R.J.;
RT "Molecular cloning of the epidermal growth factor-stimulated protein
RT kinase p56 KKIAMRE."
RL Oncogene 13:2563-2574(1996).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U35146; AAC50918.1; -.
DR HSSP; P24941; IHCL.
DR Genew; HGNC:1782; CDKL2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SMO0220; S TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 493 AA; 56318 MW; SCC30A91CBF89EFE CRC64;

Query Match 56.6%; Score 941.5; DB 4; Length 493;
Best Local Similarity 53.1%; Pred. No. 6.6e-78;
Matches 172; Conservative 68; Mismatches 69; Indels 15; Gaps 5;

CY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
CY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
CY 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
DB 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
CY 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
DB 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
CY 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
DB 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
CY 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
DB 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
CY 180 SVDIWAIGCVFAELLTGQPLWPKGSVDQYLIIRTLGKLIIRHOSIFKSNCFEGHISI 239
DB 180 SVDIWAIGCVFAELLTGQPLWPKGSVDQYLIIRTLGKLIIRHOSIFKSNCFEGHISI 239
CY 181 KAVDWAIGCLVTEMLNGEPLFFGDSIDQYLIIMRCLGNLIPRHQELFYKNPVFAGVRL 240
DB 181 KAVDWAIGCLVTEMLNGEPLFFGDSIDQYLIIMRCLGNLIPRHQELFYKNPVFAGVRL 240
CY 240 PEPEDMETLBEKFDVHPVALNFMKGLKNPDDRLTCSQLLESSYF--DSF-----QEA 292
DB 240 PEPEDMETLBEKFDVHPVALNFMKGLKNPDDRLTCSQLLESSYF--DSF-----QEA 292
CY 241 PEIKSEPLERRYPKLVSEVVIDLAKKCLHVDPKRPFCAELLHDDFFQMDGFAERFSQEL 300
DB 241 PEIKSEPLERRYPKLVSEVVIDLAKKCLHVDPKRPFCAELLHDDFFQMDGFAERFSQEL 300
CY 293 Q-K--RKARN-----EGNRPRQO 309
DB 301 QMKYQKDARNISLSKKSQNRKKEK 324

RESULT 4
Q9TTK0 PRELIMINARY; PRT; 566 AA.
ID Q9TTK0
AC Q9TTK0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KKIAMRE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA Gomi H., Sun W., Finch C.E., Itohara S., Yoshitani K., Thompson R.F.;
RT "Learning induces a cdc2-related protein kinase, KKIAMRE."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029045; BAA88508.1; -.
DR HSSP; P24941; IHCL.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SMO0220; S TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 566 AA; 64052 MW; 5D575955509C2EA9 CRC64;

Query Match 58.2%; Score 967.5; DB 6; Length 566;
Best Local Similarity 54.8%; Pred. No. 3.3e-80;
Matches 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

CY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
CY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
CY 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
DB 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
CY 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
DB 61 PNLVNLIEVFRKKRKHVLFVFCYDHTLLNELEARNPNAGVADGVKSVLWOTLQALNFCCHI 120
CY 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
DB 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
CY 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
DB 121 NCHRDIKPENILITKOGIHKICDFEFAQIL--PGDAYTDYVATRWYRAPELLVGDTOYG 179
CY 180 SVDIWAIGCVFAELLTGQPLWPKGSVDQYLIIRTLGKLIIRHOSIFKSNCFEGHISI 239
DB 180 SVDIWAIGCVFAELLTGQPLWPKGSVDQYLIIRTLGKLIIRHOSIFKSNCFEGHISI 239
CY 181 KAVDWAIGCLVTEMLNGEPLFFGDSIDQYLIIMRCLGNLIPRHQELFYKNPVFAGVRL 240
DB 181 KAVDWAIGCLVTEMLNGEPLFFGDSIDQYLIIMRCLGNLIPRHQELFYKNPVFAGVRL 240
CY 240 PEPEDMETLBEKFDVHPVALNFMKGLKNPDDRLTCSQLLESSYF--DSF-----QEA 292
DB 240 PEPEDMETLBEKFDVHPVALNFMKGLKNPDDRLTCSQLLESSYF--DSF-----QEA 292
CY 241 PEIKSEPLERRYPKLVSEVVIDLAKKCLHVDPKRPFCAELLHDDFFQMDGFAERFSQEL 300
DB 241 PEIKSEPLERRYPKLVSEVVIDLAKKCLHVDPKRPFCAELLHDDFFQMDGFAERFSQEL 300
CY 293 Q-K--RKARN-----EGNRPRQO 309
DB 301 QMKYQKDARNISLSKKSQNRKKEK 324

```

```
Db 121 NIHRDIKXPNILVSGVWKLCDGFGARTLAAPGEVTDYVATRWYRAPELLVGDVYK 180
QY 180 SSVDTWAIGCVFAELLTQGPLWPKSGVDQLYLIRTLGKLIIPRHQIFKSGFFHGISI 239
Db 181 KAVDWAIGCVFAELLTQGPLWPKSGVDQLYLIRTLGKLIIPRHQIFKSGFFHGISI 240
QY 240 PEPEMETLEEKFSVDVHPVAFNFKGCLKXNPDRDLTCSQLLESSYF--DSF-----QEA 292
Db 241 PEKDAEAPLESRYPKLSEWIDLAKKCLHIDPDRPFCAELLHDFQMDGFAERFSQEL 300
QY 293 QIK--RKARN-----EGRNRRQ 309
Db 301 QLVQKQDARNVSLSKSQNRKKEK 324

RESULT 6
ID Q9QY12 PRELIMINARY; PRT; 564 AA.
AC Q9QY12;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KXIARE-gamma.
GN CDKL2 OR KXK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KXIARE gene: variants, dual promoters, expression and
RT chromosomal localization.";
RL Submitted (JUN-1999) TO THE EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029066; BAA88428.1; -.
DR EMBL; AB029066; BAA88428.1; JOINED.
DR EMBL; AB029066; BAA88428.1; JOINED.
DR EMBL; AB029066; BAA88428.1; JOINED.
DR EMBL; AB029066; BAA88428.1; JOINED.
DR EMBL; AB029066; BAA88428.1; JOINED.
DR HSSP; P24941; IHCU.
DR MGD; MGI:185827; Cdkl2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 564 AA; 63640 MW; DACDEF630CCEB2D5 CRC64;

Query Match 56.0%; Score 930.5; DB 11; Length 564;
Best Local Similarity 53.4%; Pred. No. 8.1e-77;
Matches 174; Conservative 66; Mismatches 69; Indels 17; Gaps 6;

QY 1 MEKYELAKTSGSVGVVFKCRNKTSGQVAVKPFVESEDPPVVKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGSYGMVWCKRNDGSRIVAIFKFLSDDDKMKVKKIAMREIKLLKQLRH 60
QY 61 PNLVNLIEVFRKRWHLVFEVCCHTLNELERNPNGVADGVKSVLWQTLOALNFCIH 120
Db 61 ENLVNLLEVCCKKRWYLVFEVDHTILDLLKLPNGLDYQVQKYLFIINGIGFCHSH 120
QY 121 NCITHRIDKPNILVSGVWKLCDGFGARTLAAPGEVTDYVATRWYRAPELLVGDVYK 179
Db 121 NIHRDIKXPNILVSGVWKLCDGFGARTLAAPGEVTDYVATRWYRAPELLVGDVYK 180
QY 180 SSVDTWAIGCVFAELLTQGPLWPKSGVDQLYLIRTLGKLIIPRHQIFKSGFFHGISI 239
Db 181 KAVDWAIGCVFAELLTQGPLWPKSGVDQLYLIRTLGKLIIPRHQIFKSGFFHGISI 240
QY 240 PEPEMETLEEKFSVDVHPVAFNFKGCLKXNPDRDLTCSQLLESSYF--DSF-----Q 290
Db 241 PEKDAEAPLESRYPKLSEWIDLAKKCLHIDPDRPFCAELLHDFQMDGFAERFSQ 300
```

```
QY 291 EAQIK--RKARN-----EGRNRRQ 309
Db 301 ELQLIKEDARNVSLPKSQNRKKEK 326

RESULT 7
ID Q9QUK3 PRELIMINARY; PRT; 568 AA.
AC Q9QUK3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KXIARE.
GN CDKL2 OR KKM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; and C57BL/6;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KXIARE gene: variants, dual promoters, expression and
RT chromosomal localization.";
RL Submitted (JUN-1999) TO THE EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; ABC29073; BAA88439.1; -.
DR EMBL; ABC29069; BAA88439.1; JOINED.
DR EMBL; ABC29070; BAA88439.1; JOINED.
DR EMBL; ABC29071; BAA88439.1; JOINED.
DR EMBL; ABC29072; BAA88439.1; JOINED.
DR EMBL; ABC29065; BAA88427.1; -.
DR HSSP; P24941; IHCU.
DR YGD; MGI:185827; Cdkl2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 568 AA; 64055 MW; A43B75E2B9E86C4 CRC64;

Query Match 56.0%; Score 930.5; DB 11; Length 568;
Best Local Similarity 53.4%; Pred. No. 8.2e-77;
Matches 174; Conservative 66; Mismatches 69; Indels 17; Gaps 6;

QY 1 MEKYELAKTSGSVGVVFKCRNKTSGQVAVKPFVESEDPPVVKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGSYGMVWCKRNDGSRIVAIFKFLSDDDKMKVKKIAMREIKLLKQLRH 60
QY 61 PNLVNLIEVFRKRWHLVFEVCCHTLNELERNPNGVADGVKSVLWQTLOALNFCIH 120
Db 61 ENLVNLLEVCCKKRWYLVFEVDHTILDLLKLPNGLDYQVQKYLFIINGIGFCHSH 120
QY 121 NCITHRIDKPNILVSGVWKLCDGFGARTLAAPGEVTDYVATRWYRAPELLVGDVYK 179
Db 121 NIHRDIKXPNILVSGVWKLCDGFGARTLAAPGEVTDYVATRWYRAPELLVGDVYK 180
QY 180 SSVDTWAIGCVFAELLTQGPLWPKSGVDQLYLIRTLGKLIIPRHQIFKSGFFHGISI 239
Db 181 KAVDWAIGCVFAELLTQGPLWPKSGVDQLYLIRTLGKLIIPRHQIFKSGFFHGISI 240
QY 240 PEPEMETLEEKFSVDVHPVAFNFKGCLKXNPDRDLTCSQLLESSYF--DSF-----Q 290
Db 241 PEKDAEAPLESRYPKLSEWIDLAKKCLHIDPDRPFCAELLHDFQMDGFAERFSQ 300
QY 291 EAQIK--RKARN-----EGRNRRQ 309
```

Db	301	ELQKIEKDARNNSLPKXQNRKKEK	326	
RESULT 8				
Q9QYI1		PRELIMINARY;	PRT;	329 AA.
AC	Q9QYI1			
DT	01-MAR-2003	(Tremblrel. 13, Created)		
DT	01-MAR-2003	(Tremblrel. 13, Last sequence update)		
DE	01-MAR-2003	(Tremblrel. 23, Last annotation update)		
DE	Ser/Thr kinase KKIAMBRE-delta (Fragment).			
GN	CDK2 OR KXV.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6; TISSUE=Brain;			
RA	Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Ichihara S.;			
RT	"The Murine KKIAMBRE gene: variants, dual promoters, expression; and			
RT	chromosomal localization."			
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
DR	EMBL; AB029067; BAA88429.1; --			
DR	HSSP; P24941; IHCL.			
DR	MGI; MGI:1658227; Gdcl2.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR002290; Ser_thr_pkinase.			
DR	Pfam; PF00069; pkinase; 1.			
DR	ProDom; PD000001; Prot_kinase; 1.			
DR	SMART; SM00220; S_TKc; 1.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.			
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.			
FT	NON TER			
FT	329			
SQ	SEQUENCE	329 AA; 38023 MW; F72EB89E65AE8585 CRC64;		
Query Watch				
Best Local Similarity 55.8%; Score 928; DB 11; Length 329;				
Matches 173; Conservative 63; Mismatches 73; Indels 12; Gaps 5;				
QY	1	MEKYELAKTGBGSGVGVFKCRNKTSGGVAVKVFESDDPVVKKIALREIRMLKOLKH	60	
DB	1	MEKYENLGLVGGSGVGMVKCRNKTSGRAVAIKKFLSEDDKWKXINREIKLKQLRH	60	
QY	61	PNLVNLEVEFRKRKCHLVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCIH	120	
DB	61	ENLVNLEVEFRKRKCHLVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCIH	120	
QY	121	NCIHRDIKPENILITKOGIIKTCDFGFAQIL-IPGDAYTCVATRYAPAPELLNGCTQYG	179	
DB	121	NCIHRDIKPENILITKOGIIKTCDFGFAQIL-IPGDAYTCVATRYAPAPELLNGCTQYG	179	
QY	180	SSVDIWAIGCVFAELLTGQPLWPKSGSDVDQYLIIRTLGKLIIPRHQIFKSNFFHGISI	239	
DB	180	SPVDIWAIGCVFAELLTGQPLWPKSGSDVDQYLIIRTLGKLIIPRHQIFKSNFFHGISI	239	
QY	240	PEPEMETLEEKTSVHPVAFNFMKGLKMPDRLTCSCLLESSYF--DSFOE---AQI	294	
DB	240	PCVQHPKNARKKYPKLNGLLADIVHACLOIDPADRISSDILLHHEYFTRGCFIEKFMPEL	299	
QY	295	KRKARNEGR	303	
DB	300	KAKLLOEAK	308	
RESULT 9				
Q9P114		PRELIMINARY;	PRT;	455 AA.
AC	Q9P114			
DT	01-OCT-2000	(Tremblrel. 15, Created)		

```
RJ Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL: BC041799; AAK1799.1; -.
KW Kinase; Cyclin.
SQ SEQUENCE 592 AA; 67513 MW; 2B1AF08906E9B7697 CRC64;

Query Match 47.2%; Score 784.5; DB 4; Length 592;
Best Local Similarity 48.5%; Pred. No. 2.2e-63;
Matches 150; Conservative 61; Mismatches 91; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGOVAVKXFFVESEDDPVVYKIALREIRMLKOLKH 60
DB 1 MEMYETLCKVGEISYGVVFKCRNKTSGOVAVKXFFVESEDDPVVYKIALREIRMLKOLKH 59
61 PNLVNLIIEVFRKRKMHVFEYCDHTLLNELRPNPGVADGVIVKSLVQTLQALNFCFHH 120
DB 60 ENLVNLIIEVFRKRKMHVFEYCDHTLLNELRPNPGVADGVIVKSLVQTLQALNFCFHH 119
121 NCIHRIKIPENILITKQGIKICDFGFAQIL-IPGDATYDVATRWYRAPELLVGDGTQYG 179
DB 120 NIIHRDIKIPENILVSQSGITKLCDFGFAQTLPACDIYDVATRWYRAPELLVGDTSYG 179
180 SSVDIWAIGCVFAELLTCQPLWPGKSDVDQLYLIIRTLGKLIPIRHQSFKSNFFHGISI 239
DB 180 KPDVIALGCMIIENATGNPFLPSSDDLHLKVLKVNLSPLCNIFSKSIFAGVWL 239
240 PEPEDMETLEEFKFSVHPVALFMKGLKMPDRLTCSQLLESSYF--DSFOE---AQI 294
DB 240 PQVQPKNARKYPKLNGLLADIVRAQLQIDPACRISDDLHHEFTRGDFTEKFPPEL 295
295 KKKARNEGR 303
DB 300 KAKJLQEAR 308

RESULT 11
Q9BNG2 PRELIMINARY; PRT: 657 AA.
AC Q9BNG2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JAN-2003 (TREMBLrel. 23, Last annotation update)
DE Putative MAP kinase.
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
SEQUENCE FROM N.A.
RA Sarkar M., Matthews K.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF326965; AAG49589.1; -.
DR HSSP: P24941; I939.
DR InterPro: IPR001092; HLH basic.
DR InterPro: IPR000719; Prot_kinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00038; HLH 1; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 657 AA; 72116 MW; C93FA3C2D08F0619 CRC64;

Query Match 47.0%; Score 781; DB 5; Length 657;
Best Local Similarity 48.3%; Pred. No. 5.3e-63;
Matches 138; Conservative 67; Mismatches 81; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGOVAVKXFFVESEDDPVVYKIALREIRMLKOLKH 60
DB 1 MDAYETLGLGEGTYGVVFKARHRTSR:VAIKYKQAECDJHVRKTSLEVRVYLKQLRH 60
```

```

Db 240 PQVQHPKTKARKYVPKLNG:LA DIVHACLOIDPAERTSTSDLLRHDFYFTRDGFIEKKWTET 299
QY 294 -----MKRKARNEGRNRRQ 309
Db 300 KGPRSSKSGESLRPKGAKENSQTRRSQ 325

RESULT 13
Q9GRT9 PRELIMINARY; PRT; 1106 AA.
AC Q9GRT9;
DT 01-MAR-2003 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative mitogen-activated protein kinase 6.
GN MPK6.
OS Leishmania mexicana.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MNYC/BZ/62/M379;
RX Wiese M., Wang Q., Goercke I.;
RA "Identification of mitogen-activated protein (MAP) kinases from
RT Leishmania mexicana.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASSES.
DR EMBL: AJ293284; CAC07960.1; -.
DR HSSP: P24941; 1B38.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser thr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc.1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferrase.
SQ SEQUENCE 1106 AA; 118884 MW; 24AC26F5C209923E CRC64;

Query Match 46.3%; Score 769.5; DB 5; Length 1106;
Best Local Similarity 48.3%; Pred. No. 1.2e-61;
Matches 145; Conservative 62; Mismatches 88; Indels 5; Gaps 2;

QY 1 MEKYEKLAKTGGSGYGVFKCRNKTSGOVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEAYETLGLGEGTGVVVVKASRVTKLVA:KRFQTEQDEHVRTSSRVKMLQLLQH 60

QY 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNEJERNPNGVADGVKSVLQTLQALNFCIH 120
Db 61 PNVIRLEDVFRREGKLYLVFEFDHTILQLLESTTRGFHRELRYTYCLLRGISFCNQ 120

QY 121 NCIHRIKIPENILITKQGIKICDGFQAQLIPGDAYTDYVATRYRAPELVGDTQYGS 180
Db 121 NTHRDVKPENVLIDESGLLKLCDGFQAQTSAGKQYTDYVATRYRAPELVGVAYGK 180

QY 181 SYDVAIGCVFAELLTGQPLWPKGSVDVCLYIIRTLGKLI PRHOSIFKNSGFFREGISIP 240
Db 181 PVDVWALGCMFAELSDGQPLFGESDQQLCMTQCGVPQRLVPIFMENPLNGISFP 240

QY 241 EPEDMETLEKFSVDHPVAFNMFKMCKLKNPDRITCSQLLESSYF--DSFQ---EAQIK 295
Db 241 HTDILYTKDRYHRESNDWIEFLSSCLHTDPAQLTCTELMELPYFTRDGFRCRYEAEIR 300

RESULT 14
Q8BKX2 PRELIMINARY; PRT; 457 AA.
AC Q8BKX2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serine/threonine kinase NKIATRE alpha homolog.

Os Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX NEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL: AK050990; BAC34488.1; -. 67CDD36F51152B61 CRC64;
SQ SEQUENCE 457 AA; 51900 MW; 67CDD36F51152B61 CRC64;

Query Match 46.2%; Score 767.5; DB 11; Length 457;
Best Local Similarity 47.2%; Pred. No. 5.7e-62;
Matches 146; Conservative 64; Mismatches 92; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGGSGYGVFKCRNKTSGOVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEAYETLGLGEGTGVVVVKASRVTKLVA:KRFQTEQDEHVRTSSRVKMLQLLQH 60

QY 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNEJERNPNGVADGVKSVLQTLQALNFCIH 120
Db 61 PNVIRLEDVFRREGKLYLVFEFDHTILQLLESTTRGFHRELRYTYCLLRGISFCNQ 120

QY 121 NCIHRIKIPENILITKQGIKICDGFQAQLIPGDAYTDYVATRYRAPELVGDTQYGS 179
Db 121 NTHRDVKPENVLIDESGLLKLCDGFQAQTSAGKQYTDYVATRYRAPELVGVAYGK 179

QY 180 KPVDIWAIGCVFAELLTGQPLWPKGSVDVCLYIIRTLGKLI PRHOSIFKNSGFFREGISIP 239
Db 180 KPVDIWAIGCVFAELLTGQPLWPKGSVDVCLYIIRTLGKLI PRHOSIFKNSGFFREGISIP 239

QY 240 PEPEMETLEKFSVDHPVAFNMFKMCKLKNPDRITCSQLLESSYF--DSFQ---AQI 294
Db 240 PQVQHPKTKARKYVPKLNG:LA DIVHACLOIDPAERTSTSDLLRHDFYFTRDGFIEKKWTET 299

QY 295 KRKARNEGR 303
Db 300 RAKLQEAQ 308

RESULT 15
Q8BL49 PRELIMINARY; PRT; 585 AA.
AC Q8BL49;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serine/threonine kinase NKIATRE alpha homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL: AK046394; BAC32701.1; -.
SQ SEQUENCE 585 AA; 66530 MW; A1867A2A6CDF15A2 CRC64;

Query Match 46.0%; Score 764.5; DB 11; Length 585;
Best Local Similarity 46.9%; Pred. No. 1.5e-61;
Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;

```

